2AC

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
6 March 2003 (06.03.2003)

PCT

(10) International Publication Number
WO 03/018812 A1

(51) International Patent Classification[7]: C12N 15/52, 9/00, 9/02, C07K 14/37, C12N 15/82, 1/21, 15/53, 15/31, G01N 33/573, 33/68, A61K 35/00, A01N 61/00, A01H 5/00

(21) International Application Number: PCT/EP02/09331

(22) International Filing Date: 21 August 2002 (21.08.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/313,810 22 August 2001 (22.08.2001) US

(71) Applicant *(for all designated States except US)*: **BAYER AKTIENGESELLSCHAFT** [DE/DE]; 51368 Leverkusen (DE).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: **SMOLYAR, Alex** [US/US]; 734 Boylston Street, Brookline, MA 02467 (US).

(74) Common Representative: **BAYER AKTIENGESELLSCHAFT**; 51368 Leverkusen (DE).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)*

— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for all designations*

Published:

— *with international search report*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 03/018812 A1

(54) Title: REGULATION OF HUMAN PTPASE-LIKE PROTEIN

(57) Abstract: Reagents that regulate human PTPase-like protein and reagents which bind to human PTPase-like protein gene products can play a role in preventing, ameliorating, or correcting dysfunctions or diseases including, but not limited to, obesity, diabetes, CNS disorders, cardiovascular disorders or COPD.

REGULATION OF HUMAN PTPASE-LIKE PROTIEN

TECHNICAL FIELD OF THE INVENTION

The invention relates to the regulation of human PTPase-like protein.

BACKGROUND OF THE INVENTION

Phosphorylation of proteins is a fundamental mechanism for regulating diverse cellular processes. U.S. Patent 5,952,212. While the majority of protein phosphorylation occurs at serine and threonine residues, phosphorylation at tyrosine residues is attracting a great deal of interest since the discovery that many oncogene products and growth factor receptors possess intrinsic protein tyrosine kinase activity. The importance of protein tyrosine phosphorylation in growth factor signal transduction, cell cycle progression and neoplastic transformation is now well established. Hunter *et al.*, *Ann. Rev. Biochem. 54*, 987-30, 1985; Ullrich *et al.*, Cell 61, 203-12, 1990; Nurse, *Nature 344*, 503-08, 1990; Cantley *et al.*, *Cell 64*, 281-302, 1991. Biochemical studies have shown that phosphorylation on tyrosine residues of a variety of cellular proteins is a dynamic process involving competing phosphorylation and dephosphorylation reactions. The regulation of protein tyrosine phosphorylation is mediated by the reciprocal actions of protein tyrosine kinases (PTKases or PTKS) and protein tyrosine phosphatases (PTPs). The tyrosine phosphorylation reactions are catalyzed by PTKs. Tyrosine phosphorylated proteins can be specifically dephosphorylated through the action of PTPs. The level of protein tyrosine phosphorylation of intracellular substances is determined by the balance of PTK and PTP activities. Hunter, *Cell 58*, 1013-16, 1989.

Protein Tyrosine Kinases

The protein tyrosine kinases (PTKS; ATP:protein-tyrosine O-phosphotransferase, EC 2.7.1.112) are a large family of proteins that includes many growth factor receptors and potential oncogenes. Hanks *et al.*, *Science 241*, 42-52, 1988. Many PTKs have

been linked to initial signals required for induction of the cell cycle. Weaver *et al.*, *Mol. Cell. Biol. 11(9)*, 4415-22, 1991. PTKs comprise a discrete family of enzymes having common ancestry with, but major differences from, serine/threonine-specific protein kinases. Hanks *et al.*, *supra*. The mechanisms leading to changes in activity of PTKs are best understood in the case of receptor-type PTKs having a transmembrane topology. Ullrich *et al.*, 1990, *supra*. The binding of specific ligands to the extracellular domain of members of receptor-type PTKs is thought to induce their oligomerization, leading to an increase in tyrosine kinase activity and activation of the signal transduction pathways. Ullrich *et al.*, 1990, *supra*. Deregulation of kinase activity through mutation or overexpression is a well-established mechanism for cell transformation. Hunter *et al.*, 1985; Ullrich *et al.*, 1990.

Protein Tyrosine Phosphatases

The protein phosphatases are composed of at least two separate and distinct families: the protein serine/threonine phosphatases and the protein tyrosine phosphatases (PTPs; protein-tyrosine-phosphate phosphohydrolase, EC 3.13.48. Hunter, 1989. The PTPs are a family of proteins that have been classified into two subgroups. The first subgroup is made up of the low molecular weight, intracellular enzymes that contain a single conserved catalytic phosphatase domain. All known intracellular type PTPs contain a single conserved catalytic phosphatase domain. Examples of the first group of PTPs include (1) placental PTP 1B (Charbonneau *et al.*, *Proc. Natl. Acad. Sci. USA 86*, 5252-56, 1989; Chernoff *et al.*, *Proc. Natl. Acad. Sci. USA 87*, 2735-89, 1990, (2) T-cell PTP (Cool *et al.*, *Proc. Natl. Acad. Sci. USA 86*, 5257-61, 1989, (3) rat brain PTP (Guan *et al.*, *Proc. Natl. Acad. Sci. USA 87*, 1501-02, 1990, (4) neuronal phosphatase (STEP) (Lombroso *et al.*, *Proc. Natl. Acad. Sci. USA 88*, 7242-46, 1991, and (5) cytoplasmic phosphatases that contain a region of homology to cytoskeletal proteins (Gu *et al.*, *Proc. Natl. Acad. Sci. USA 88*, 5867-871, 1991; Yang *et al.*, *Proc. Natl. Acad. Sci. USA 88*, 5949-53, 1991.

The second subgroup is made up of the high molecular weight, receptor-linked PTPs, termed RPTPs. RPTPs consist of (a) an intracellular catalytic region, (b) a single

transmembrane segment, and (c) a putative ligand-binding extracellular domain. The structures and sizes of the putative ligand-binding extracellular "receptor" domains of RPTPs are quite divergent. In contrast, the intracellular catalytic regions of RPTPs are highly homologous. All RPTPs have two tandemly duplicated catalytic phosphatase homology domains, with the prominent exception of an RPTP termed HPTPβ, which has only one catalytic phosphatase domain. Tsai *et al.*, *J. Biol. Chem. 266*, 0534-43, 1991.

Since the initial purification, sequencing, and cloning of a PTP, additional potential PTPs have been identified at a rapid pace. The number of different PTPs that have been identified is increasing steadily, leading to speculations that this family may be as large as the PTK family. Hunter, 1989.

Conserved amino acid sequences designated "consensus sequences" have been identified in the catalytic domains of known PTPs. Krueger *et al.*, *EMBO J. 9*, 3241-52, 1990; Yi *et al.*, *Mol. Cell. Biol. 12*, 836-46, 1992. Yi *et al.* aligned the catalytic phosphatase domain sequences of the following PTPs: LCA, PTP1B, TCPTP, LAR, DLAR, and HPTPα, HPTPβ, and HPTPγ. This alignment includes the following consensus sequences (Yi *et al.*, *supra*):

1. K C X X Y W P (SEQ ID NO: 14)
2. H C S X G X G R X G (SEQ ID NO: 15).

Krueger et al., aligned the catalytic phosphatase domain sequences of PTP1B, TCPTP, LAR, LCA, HPTPα, β, δ, ε and ζ, and DLAR and DPTP. This alignment includes the following consensus sequences (Krueger *et al.*, *supra*):

1. K C X X Y W P
2. H C S X G X G R X G.

It is becoming clear that dephosphorylation of tyrosine residues can by itself function as an important regulatory mechanism. Dephosphorylation of a C-terminal tyrosine residue has been shown to activate tyrosine kinase activity in the src family of

tyrosine kinases. Hunter, *Cell 49*, 1-4, 1987. Tyrosine dephosphorylation has been suggested to be an obligatory step in the mitotic activation of the maturation-promoting factor (MPF) kinase. Morla *et al.*, *Cell 58*, 193-203, 1989. These observations point out the need in the art for understanding the mechanisms that regulate tyrosine phosphatase activity.

Because of the importance of these enzymes in a variety of cellular processes, there is a continuing need in the art to identify additional tyrosine phosphatases, which can be regulated to provide therapeutic effects.

SUMMARY OF THE INVENTION

It is an object of the invention to provide reagents and methods of regulating a human PTPase-like protein. This and other objects of the invention are provided by one or more of the embodiments described below.

One embodiment of the invention is a PTPase-like protein polypeptide comprising an amino acid sequence selected from the group consisting of:

amino acid sequences which are at least about 95% identical to the amino acid sequence shown in SEQ ID NO: 2; and

the amino acid sequence shown in SEQ ID NO: 2.

Yet another embodiment of the invention is a method of screening for agents which decrease extracellular matrix degradation. A test compound is contacted with a PTPase-like protein polypeptide comprising an amino acid sequence selected from the group consisting of:

amino acid sequences which are at least about 95% identical to the amino acid sequence shown in SEQ ID NO: 2; and

the amino acid sequence shown in SEQ ID NO: 2.
Binding between the test compound and the PTPase-like protein polypeptide is detected. A test compound which binds to the PTPase-like protein polypeptide is thereby identified as a potential agent for decreasing extracellular matrix degradation. The agent can work by decreasing the activity of the PTPase-like protein.

Another embodiment of the invention is a method of screening for agents which decrease extracellular matrix degradation. A test compound is contacted with a polynucleotide encoding a PTPase-like protein polypeptide, wherein the polynucleotide comprises a nucleotide sequence selected from the group consisting of:

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 1; and

the nucleotide sequence shown in SEQ ID NO: 1.

Binding of the test compound to the polynucleotide is detected. A test compound which binds to the polynucleotide is identified as a potential agent for decreasing extracellular matrix degradation. The agent can work by decreasing the amount of the PTPase-like protein through interacting with the PTPase-like protein mRNA.

Another embodiment of the invention is a method of screening for agents which regulate extracellular matrix degradation. A test compound is contacted with a PTPase-like protein polypeptide comprising an amino acid sequence selected from the group consisting of:

amino acid sequences which are at least about 95% identical to the amino acid sequence shown in SEQ ID NO: 2; and

the amino acid sequence shown in SEQ ID NO: 2.

A PTPase-like protein activity of the polypeptide is detected. A test compound which increases PTPase-like protein activity of the polypeptide relative to PTPase-like protein activity in the absence of the test compound is thereby identified as a potential agent for increasing extracellular matrix degradation. A test compound which decreases PTPase-like protein activity of the polypeptide relative to PTPase-like protein activity in the absence of the test compound is thereby identified as a potential agent for decreasing extracellular matrix degradation.

Even another embodiment of the invention is a method of screening for agents which decrease extracellular matrix degradation. A test compound is contacted with a PTPase-like protein product of a polynucleotide which comprises a nucleotide sequence selected from the group consisting of:

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 1; and

the nucleotide sequence shown in SEQ ID NO: 1.

Binding of the test compound to the PTPase-like protein product is detected. A test compound which binds to the PTPase-like protein product is thereby identified as a potential agent for decreasing extracellular matrix degradation.

Still another embodiment of the invention is a method of reducing extracellular matrix degradation. A cell is contacted with a reagent which specifically binds to a polynucleotide encoding a PTPase-like protein polypeptide or the product encoded by the polynucleotide, wherein the polynucleotide comprises a nucleotide sequence selected from the group consisting of:

nucleotide sequences which are at least about 50% identical to the nucleotide sequence shown in SEQ ID NO: 1; and

the nucleotide sequence shown in SEQ ID NO: 1.

PTPase-like protein activity in the cell is thereby decreased.

The invention thus provides a human PTPase-like protein that can be used to identify test compounds that may act, for example, as activators or inhibitors at the enzyme's active site. Human PTPase-like protein and fragments thereof also are useful in raising specific antibodies that can block the enzyme and effectively reduce its activity.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 shows the DNA-sequence encoding a PTPase-like protein polypeptide (SEQ ID NO: 1).

Fig. 2 shows the amino acid sequence (SEQ ID NO: 2) deduced from the DNA-sequence of Fig.1.

Fig. 3 shows the amino acid sequence of the protein identified by trembl|AB046040|AB046040_1 (SEQ ID NO:3).

Fig. 4 shows the BLASTP - alignment of 406 (SEQ ID NO: 1) against trembl|AB046040|AB046040_1 (SEQ ID NO:3).

Fig. 5 shows the HMMPFAM - alignment of 406 (SEQ ID NO: 1) against pfam|hmm|DSPc.

DETAILED DESCRIPTION OF THE INVENTION

The invention relates to an isolated polynucleotide from the group consisting of:

a) a polynucleotide encoding a PTPase-like protein polypeptide comprising an amino acid sequence selected from the group consisting of:
amino acid sequences which are at least about 95% identical to

the amino acid sequence shown in SEQ ID NO: 2; and
the amino acid sequence shown in SEQ ID NO: 2;

b) a polynucleotide comprising the sequence of SEQ ID NO: 1;

c) a polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) and (b) and encodes a PTPase-like protein polypeptide;

d) a polynucleotide the sequence of which deviates from the polynucleotide sequences specified in (a) to (c) due to the degeneration of the genetic code and encodes a PTPase-like protein polypeptide; and

e) a polynucleotide which represents a fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (d) and encodes a PTPase-like protein polypeptide.

Furthermore, it has been discovered by the present applicant that a novel PTPase-like protein, particularly a human PTPase-like protein, can be used in therapeutic methods to treat obesity, diabetes, a CNS disorder, a cardiovascular disorder or COPD. Human PTPase-like protein comprises the amino acid sequence shown in SEQ ID NO: 2. A coding sequence for human PTPase-like protein is shown in SEQ ID NO: 1. This sequence is located on chromosome 9. Related ESTs (AA907041; AI023460; BG116082; AI829572; AA910424; AA651920; BE897818; Z44764) [SEQ ID NOs: 4 to 12] are expressed in fetal liver, fetal spleen, duodenal adenocarcinoma, germinal center B cells, melanotic melanoma, and total brain.

Human PTPase-like protein is 94% identical over 682 amino acids to PTPase-like protein (FIG. 1). Human PTPase-like protein of the invention is expected to be useful for the same purposes as previously identified PTPase-like protein enzymes. Human PTPase-like protein is believed to be useful in therapeutic methods to treat disorders such as obesity, diabetes, CNS disorders, cardiovascular disorders, and COPD. Human PTPase-like protein also can be used to screen for human PTPase-like protein activators and inhibitors.

Polypeptides

Human PTPase-like protein polypeptides according to the invention comprise at least 6, 10, 15, 20, 25, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 600, 650 or 682 contiguous amino acids selected from the amino acid sequence shown in SEQ ID NO: 2 or a biologically active variant thereof, as defined below. A PTPase-like protein polypeptide of the invention therefore can be a portion of a PTPase-like protein protein, a full-length PTPase-like protein protein, or a fusion protein comprising all or a portion of a PTPase-like protein protein.

Biologically Active Variants

PTPase-like protein polypeptide variants that are biologically active, *i.e.*, retain the ability to bind a ligand to produce a biological effect, such as cyclic AMP formation, mobilization of intracellular calcium, or phosphoinositide metabolism, also are PTPase-like protein polypeptides. Preferably, naturally or non-naturally occurring PTPase-like protein polypeptide variants have amino acid sequences which are at least about 95, preferably at least about, 96, or 98% identical to the amino acid sequence shown in SEQ ID NO: 2 or a fragment thereof. Percent identity between a putative PTPase-like protein polypeptide variant and an amino acid sequence of SEQ ID NO: 2 is determined by conventional methods. See, for example, Altschul *et al., Bull. Math. Bio. 48*:603 (1986), and Henikoff & Henikoff, *Proc. Natl. Acad. Sci. USA 89*:10915 (1992). Briefly, two amino acid sequences are aligned to optimize the alignment scores using a gap opening penalty of 10, a gap extension penalty of 1, and the "BLOSUM62" scoring matrix of Henikoff & Henikoff, 1992.

Those skilled in the art appreciate that there are many established algorithms available to align two amino acid sequences. The "FASTA" similarity search algorithm of Pearson & Lipman is a suitable protein alignment method for examining the level of identity shared by an amino acid sequence disclosed herein and the amino acid sequence of a putative variant. The FASTA algorithm is described by Pearson

& Lipman, *Proc. Nat'l Acad. Sci. USA 85*:2444(1988), and by Pearson, *Meth. Enzymol. 183*:63 (1990). Briefly, FASTA first characterizes sequence similarity by identifying regions shared by the query sequence (*e.g.*, SEQ ID NO: 2) and a test sequence that have either the highest density of identities (if the ktup variable is 1) or pairs of identities (if ktup=2), without considering conservative amino acid substitutions, insertions, or deletions. The ten regions with the highest density of identities are then rescored by comparing the similarity of all paired amino acids using an amino acid substitution matrix, and the ends of the regions are "trimmed" to include only those residues that contribute to the highest score. If there are several regions with scores greater than the "cutoff" value (calculated by a predetermined formula based upon the length of the sequence the ktup value), then the trimmed initial regions are examined to determine whether the regions can be joined to form an approximate alignment with gaps. Finally, the highest scoring regions of the two amino acid sequences are aligned using a modification of the Needleman-Wunsch-Sellers algorithm (Needleman & Wunsch, *J. Mol. Biol.48*:444 (1970); Sellers, *SIAM J. Appl. Math.26*:787 (1974)), which allows for amino acid insertions and deletions. Preferred parameters for FASTA analysis are: ktup=1, gap opening penalty=10, gap extension penalty=1, and substitution matrix=BLOSUM62. These parameters can be introduced into a FASTA program by modifying the scoring matrix file ("SMATRIX"), as explained in Appendix 2 of Pearson, *Meth. Enzymol. 183*:63 (1990).

FASTA can also be used to determine the sequence identity of nucleic acid molecules using a ratio as disclosed above. For nucleotide sequence comparisons, the ktup value can range between one to six, preferably from three to six, most preferably three,. with other parameters set as default.

Variations in percent identity can be due, for example, to amino acid substitutions, insertions, or deletions. Amino acid substitutions are defined as one for one amino acid replacements. They are conservative in nature when the substituted amino acid has similar structural and/or chemical properties. Examples of conservative

replacements are substitution of a leucine with an isoleucine or valine, an aspartate with a glutamate, or a threonine with a serine.

Amino acid insertions or deletions are changes to or within an amino acid sequence. They typically fall in the range of about 1 to 5 amino acids. Guidance in determining which amino acid residues can be substituted, inserted, or deleted without abolishing biological or immunological activity of a PTPase-like protein polypeptide can be found using computer programs well known in the art, such as DNASTAR software.

The invention additionally, encompasses PTPase-like protein polypeptides that are differentially modified during or after translation, *e.g.*, by glycosylation, acetylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to an antibody molecule or other cellular ligand, etc. Any of numerous chemical modifications can be carried out by known techniques including, but not limited, to specific chemical cleavage by cyanogen bromide, trypsin, chymotrypsin, papain, V8 protease, $NaBH_4$, acetylation, formylation, oxidation, reduction, metabolic synthesis in the presence of tunicamycin, etc.

Additional post-translational modifications encompassed by the invention include, for example, *e.g.*, N-linked or O-linked carbohydrate chains, processing of N-terminal or C-terminal ends), attachment of chemical moieties to the amino acid backbone, chemical modifications of N-linked or O-linked carbohydrate chains, and addition or deletion of an N-terminal methionine residue as a result of prokaryotic host cell expression. The PTPase-like protein polypeptides may also be modified with a detectable label, such as an enzymatic, fluorescent, isotopic or affinity label to allow for detection and isolation of the protein.

The invention also provides chemically modified derivatives of PTPase-like protein polypeptides that may provide additional advantages such as increased solubility, stability and circulating time of the polypeptide, or decreased immunogenicity (see U.S. Patent No. 4,179,337). The chemical moieties for derivitization can be selected

from water soluble polymers such as polyethylene glycol, ethylene glycol/propylene glycol copolymers, carboxymethylcellulose, dextran, polyvinyl alcohol, and the like. The polypeptides can be modified at random or predetermined positions within the molecule and can include one, two, three, or more attached chemical moieties.

Whether an amino acid change or a polypeptide modification results in a biologically active PTPase-like protein polypeptide can readily be determined by assaying for binding to a ligand or by conducting a functional assay, as described for example, in Calvert-Evers & Hammond, *Cell Biol. Int. 24*, 559-68, 2000.

Fusion Proteins

Fusion proteins are useful for generating antibodies against PTPase-like protein polypeptide amino acid sequences and for use in various assay systems. For example, fusion proteins can be used to identify proteins that interact with portions of a PTPase-like protein polypeptide. Protein affinity chromatography or library-based assays for protein-protein interactions, such as the yeast two-hybrid or phage display systems, can be used for this purpose. Such methods are well known in the art and also can be used as drug screens.

A PTPase-like protein polypeptide fusion protein comprises two polypeptide segments fused together by means of a peptide bond. The first polypeptide segment comprises at least 6, 10, 15, 20, 25, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 600, or 682 contiguous amino acids of SEQ ID NO: 2 or of a biologically active variant, such as those described above. The first polypeptide segment also can comprise full-length PTPase-like protein protein.

The second polypeptide segment can be a full-length protein or a protein fragment. Proteins commonly used in fusion protein construction include β-galactosidase, β-glucuronidase, green fluorescent protein (GFP), autofluorescent proteins, including blue fluorescent protein (BFP), glutathione-S-transferase (GST), luciferase,

horseradish peroxidase (HRP), and chloramphenicol acetyltransferase (CAT). Additionally, epitope tags are used in fusion protein constructions, including histidine (His) tags, FLAG tags, influenza hemagglutinin (HA) tags, Myc tags, VSV-G tags, and thioredoxin (Trx) tags. Other fusion constructions can include maltose binding protein (MBP), S-tag, Lex a DNA binding domain (DBD) fusions, GAL4 DNA binding domain fusions, and herpes simplex virus (HSV) BP16 protein fusions. A fusion protein also can be engineered to contain a cleavage site located between the PTPase-like protein polypeptide-encoding sequence and the heterologous protein sequence, so that the PTPase-like protein polypeptide can be cleaved and purified away from the heterologous moiety.

A fusion protein can be synthesized chemically, as is known in the art. Preferably, a fusion protein is produced by covalently linking two polypeptide segments or by standard procedures in the art of molecular biology. Recombinant DNA methods can be used to prepare fusion proteins, for example, by making a DNA construct which comprises coding sequences selected from SEQ ID NO: 1 in proper reading frame with nucleotides encoding the second polypeptide segment and expressing the DNA construct in a host cell, as is known in the art. Many kits for constructing fusion proteins are available from companies such as Promega Corporation (Madison, WI), Stratagene (La Jolla, CA), CLONTECH (Mountain View, CA), Santa Cruz Biotechnology (Santa Cruz, CA), MBL International Corporation (MIC; Watertown, MA), and Quantum Biotechnologies (Montreal, Canada; 1-888-DNA-KITS).

Identification of Species Homologs

Species homologs of human PTPase-like protein polypeptide can be obtained using PTPase-like protein polypeptide polynucleotides (described below) to make suitable probes or primers for screening cDNA expression libraries from other species, such as mice, monkeys, or yeast, identifying cDNAs which encode homologs of PTPase-like protein polypeptide, and expressing the cDNAs as is known in the art.

Polynucleotides

A PTPase-like protein polynucleotide can be single- or double-stranded and comprises a coding sequence or the complement of a coding sequence for a PTPase-like protein polypeptide. A coding sequence for human PTPase-like protein is shown in SEQ ID NO: 1.

Degenerate nucleotide sequences encoding human PTPase-like protein polypeptides, as well as homologous nucleotide sequences which are at least about 50, 55, 60, 65, 70, preferably about 75, 90, 96, 98, or 99% identical to the nucleotide sequence shown in SEQ ID NO: 1 or its complement also are PTPase-like protein polynucleotides. Percent sequence identity between the sequences of two polynucleotides is determined using computer programs such as ALIGN which employ the FASTA algorithm, using an affine gap search with a gap open penalty of -12 and a gap extension penalty of -2. Complementary DNA (cDNA) molecules, species homologs, and variants of PTPase-like protein polynucleotides that encode biologically active PTPase-like protein polypeptides also are PTPase-like protein polynucleotides. Polynucleotide fragments comprising at least 8, 9, 10, 11, 12, 15, 20, or 25 contiguous nucleotides of SEQ ID NO: 1 or its complement also are PTPase-like protein polynucleotides. These fragments can be used, for example, as hybridization probes or as antisense oligonucleotides.

Identification of Polynucleotide Variants and Homologs

Variants and homologs of the PTPase-like protein polynucleotides described above also are PTPase-like protein polynucleotides. Typically, homologous PTPase-like protein polynucleotide sequences can be identified by hybridization of candidate polynucleotides to known PTPase-like protein polynucleotides under stringent conditions, as is known in the art. For example, using the following wash

conditions--2X SSC (0.3 M NaCl, 0.03 M sodium citrate, pH 7.0), 0.1% SDS, room temperature twice, 30 minutes each; then 2X SSC, 0.1% SDS, 50 °C once, 30 minutes; then 2X SSC, room temperature twice, 10 minutes each--homologous sequences can be identified which contain at most about 25-30% basepair mismatches. More preferably, homologous nucleic acid strands contain 15-25% basepair mismatches, even more preferably 5-15% basepair mismatches.

Species homologs of the PTPase-like protein polynucleotides disclosed herein also can be identified by making suitable probes or primers and screening cDNA expression libraries from other species, such as mice, monkeys, or yeast. Human variants of PTPase-like protein polynucleotides can be identified, for example, by screening human cDNA expression libraries. It is well known that the T_m of a double-stranded DNA decreases by 1-1.5 °C with every 1% decrease in homology (Bonner *et al.*, *J. Mol. Biol. 81*, 123 (1973). Variants of human PTPase-like protein polynucleotides or PTPase-like protein polynucleotides of other species can therefore be identified by hybridizing a putative homologous PTPase-like protein polynucleotide with a polynucleotide having a nucleotide sequence of SEQ ID NO: 1 or the complement thereof to form a test hybrid. The melting temperature of the test hybrid is compared with the melting temperature of a hybrid comprising polynucleotides having perfectly complementary nucleotide sequences, and the number or percent of basepair mismatches within the test hybrid is calculated.

Nucleotide sequences which hybridize to PTPase-like protein polynucleotides or their complements following stringent hybridization and/or wash conditions also are PTPase-like protein polynucleotides. Stringent wash conditions are well known and understood in the art and are disclosed, for example, in Sambrook *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL, 2d ed., 1989, at pages 9.50-9.51.

Typically, for stringent hybridization conditions a combination of temperature and salt concentration should be chosen that is approximately 12-20 °C below the calculated T_m of the hybrid under study. The T_m of a hybrid between a PTPase-like

protein polynucleotide having a nucleotide sequence shown in SEQ ID NO: 1 or the complement thereof and a polynucleotide sequence which is at least about 50, preferably about 75, 90, 96, or 98% identical to one of those nucleotide sequences can be calculated, for example, using the equation of Bolton and McCarthy, *Proc. Natl. Acad. Sci. U.S.A. 48*, 1390 (1962):

$$T_m = 81.5\ ^{\circ}C - 16.6(\log_{10}[Na^+]) + 0.41(\%G + C) - 0.63(\%\text{formamide}) - 600/l),$$

where l = the length of the hybrid in basepairs.

Stringent wash conditions include, for example, 4X SSC at 65°C, or 50% formamide, 4X SSC at 42°C, or 0.5X SSC, 0.1% SDS at 65°C. Highly stringent wash conditions include, for example, 0.2X SSC at 65°C.

Preparation of Polynucleotides

A PTPase-like protein polynucleotide can be isolated free of other cellular components such as membrane components, proteins, and lipids. Polynucleotides can be made by a cell and isolated using standard nucleic acid purification techniques, or synthesized using an amplification technique, such as the polymerase chain reaction (PCR), or by using an automatic synthesizer. Methods for isolating polynucleotides are routine and are known in the art. Any such technique for obtaining a polynucleotide can be used to obtain isolated PTPase-like protein polynucleotides. For example, restriction enzymes and probes can be used to isolate polynucleotide fragments, which comprise PTPase-like protein nucleotide sequences. Isolated polynucleotides are in preparations that are free or at least 70, 80, or 90% free of other molecules.

Human PTPase-like protein cDNA molecules can be made with standard molecular biology techniques, using PTPase-like protein mRNA as a template. Human PTPase-like protein cDNA molecules can thereafter be replicated using molecular biology techniques known in the art and disclosed in manuals such as Sambrook *et al.* (1989).

An amplification technique, such as PCR, can be used to obtain additional copies of polynucleotides of the invention, using either human genomic DNA or cDNA as a template.

Alternatively, synthetic chemistry techniques can be used to synthesize PTPase-like protein polynucleotides. The degeneracy of the genetic code allows alternate nucleotide sequences to be synthesized which will encode a PTPase-like protein polypeptide having, for example, an amino acid sequence shown in SEQ ID NO: 2 or a biologically active variant thereof.

Extending Polynucleotides

Various PCR-based methods can be used to extend the nucleic acid sequences disclosed herein to detect upstream sequences such as promoters and regulatory elements. For example, restriction-site PCR uses universal primers to retrieve unknown sequence adjacent to a known locus (Sarkar, *PCR Methods Applic. 2*, 318-322, 1993). Genomic DNA is first amplified in the presence of a primer to a linker sequence and a primer specific to the known region. The amplified sequences are then subjected to a second round of PCR with the same linker primer and another specific primer internal to the first one. Products of each round of PCR are transcribed with an appropriate RNA polymerase and sequenced using reverse transcriptase.

Inverse PCR also can be used to amplify or extend sequences using divergent primers based on a known region (Triglia *et al.*, *Nucleic Acids Res. 16*, 8186, 1988). Primers can be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences Inc., Plymouth, Minn.), to be 22-30 nucleotides in length, to have a GC content of 50% or more, and to anneal to the target sequence at temperatures about 68-72 °C. The method uses several restriction enzymes to generate a suitable fragment in the known region of a gene. The fragment is then circularized by intramolecular ligation and used as a PCR template.

Another method which can be used is capture PCR, which involves PCR amplification of DNA fragments adjacent to a known sequence in human and yeast artificial chromosome DNA (Lagerstrom *et al.*, *PCR Methods Applic. 1*, 111-119, 1991). In this method, multiple restriction enzyme digestions and ligations also can be used to place an engineered double-stranded sequence into an unknown fragment of the DNA molecule before performing PCR.

Another method which can be used to retrieve unknown sequences is that of Parker *et al.*, *Nucleic Acids Res. 19*, 3055-3060, 1991). Additionally, PCR, nested primers, and PROMOTERFINDER libraries (CLONTECH, Palo Alto, Calif.) can be used to walk genomic DNA (CLONTECH, Palo Alto, Calif.). This process avoids the need to screen libraries and is useful in finding intron/exon junctions.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. Randomly-primed libraries are preferable, in that they will contain more sequences which contain the 5' regions of genes. Use of a randomly primed library may be especially preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries can be useful for extension of sequence into 5' non-transcribed regulatory regions.

Commercially available capillary electrophoresis systems can be used to analyze the size or confirm the nucleotide sequence of PCR or sequencing products. For example, capillary sequencing can employ flowable polymers for electrophoretic separation, four different fluorescent dyes (one for each nucleotide) that are laser activated, and detection of the emitted wavelengths by a charge coupled device camera. Output/light intensity can be converted to electrical signal using appropriate software (*e.g.* GENOTYPER and Sequence NAVIGATOR, Perkin Elmer), and the entire process from loading of samples to computer analysis and electronic data display can be computer controlled. Capillary electrophoresis is especially preferable

for the sequencing of small pieces of DNA that might be present in limited amounts in a particular sample.

Obtaining Polypeptides

Human PTPase-like protein polypeptides can be obtained, for example, by purification from human cells, by expression of PTPase-like protein polynucleotides, or by direct chemical synthesis.

Protein Purification

Human PTPase-like protein polypeptides can be purified from any cell that expresses the polypeptide, including host cells that have been transfected with PTPase-like protein expression constructs. A purified PTPase-like protein polypeptide is separated from other compounds that normally associate with the PTPase-like protein polypeptide in the cell, such as certain proteins, carbohydrates, or lipids, using methods well-known in the art. Such methods include, but are not limited to, size exclusion chromatography, ammonium sulfate fractionation, ion exchange chromatography, affinity chromatography, and preparative gel electrophoresis. A preparation of purified PTPase-like protein polypeptides is at least 80% pure; preferably, the preparations are 90%, 95%, or 99% pure. Purity of the preparations can be assessed by any means known in the art, such as SDS-polyacrylamide gel electrophoresis.

Expression of Polynucleotides

To express a PTPase-like protein polynucleotide, the polynucleotide can be inserted into an expression vector that contains the necessary elements for the transcription and translation of the inserted coding sequence. Methods that are well known to those skilled in the art can be used to construct expression vectors containing sequences encoding PTPase-like protein polypeptides and appropriate transcriptional

and translational control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. Such techniques are described, for example, in Sambrook *et al.* (1989) and in Ausubel *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, N.Y., 1989.

A variety of expression vector/host systems can be utilized to contain and express sequences encoding a PTPase-like protein polypeptide. These include, but are not limited to, microorganisms, such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors, insect cell systems infected with virus expression vectors (*e.g.*, baculovirus), plant cell systems transformed with virus expression vectors (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (*e.g.*, Ti or pBR322 plasmids), or animal cell systems.

The control elements or regulatory sequences are those non-translated regions of the vector -- enhancers, promoters, 5' and 3' untranslated regions -- which interact with host cellular proteins to carry out transcription and translation. Such elements can vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, can be used. For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the BLUESCRIPT phagemid (Stratagene, LaJolla, Calif.) or pSPORT1 plasmid (Life Technologies) and the like can be used. The baculovirus polyhedrin promoter can be used in insect cells. Promoters or enhancers derived from the genomes of plant cells (*e.g.*, heat shock, RUBISCO, and storage protein genes) or from plant viruses (*e.g.*, viral promoters or leader sequences) can be cloned into the vector. In mammalian cell systems, promoters from mammalian genes or from mammalian viruses are preferable. If it is necessary to generate a cell line that contains multiple copies of a nucleotide sequence encoding a PTPase-like protein polypeptide, vectors based on SV40 or EBV can be used with an appropriate selectable marker.

Bacterial and Yeast Expression Systems

In bacterial systems, a number of expression vectors can be selected depending upon the use intended for the PTPase-like protein polypeptide. For example, when a large quantity of a PTPase-like protein polypeptide is needed for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified can be used. Such vectors include, but are not limited to, multifunctional *E. coli* cloning and expression vectors such as BLUESCRIPT (Stratagene). In a BLUESCRIPT vector, a sequence encoding the PTPase-like protein polypeptide can be ligated into the vector in frame with sequences for the amino-terminal Met and the subsequent 7 residues of β-galactosidase so that a hybrid protein is produced. pIN vectors (Van Heeke & Schuster, *J. Biol. Chem. 264*, 5503-5509, 1989) or pGEX vectors (Promega, Madison, Wis.) also can be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems can be designed to include heparin, thrombin, or factor Xa protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

In the yeast *Saccharomyces cerevisiae*, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH can be used. For reviews, see Ausubel *et al.* (1989) and Grant *et al.*, *Methods Enzymol. 153*, 516-544, 1987.

Plant and Insect Expression Systems

If plant expression vectors are used, the expression of sequences encoding PTPase-like protein polypeptides can be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of CaMV can be used

alone or in combination with the omega leader sequence from TMV (Takamatsu, *EMBO J. 6*, 307-311, 1987). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters can be used (Coruzzi *et al.*, *EMBO J. 3*, 1671-1680, 1984; Broglie *et al.*, *Science 224*, 838-843, 1984; Winter *et al.*, *Results Probl. Cell Differ. 17*, 85-105, 1991). These constructs can be introduced into plant cells by direct DNA transformation or by pathogen-mediated transfection. Such techniques are described in a number of generally available reviews (*e.g.*, Hobbs or Murray, in McGraw Hill Yearbook of Science and Technology, McGraw Hill, New York, N.Y., pp. 191-196, 1992).

An insect system also can be used to express a PTPase-like protein polypeptide. For example, in one such system *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in *Spodoptera frugiperda* cells or in *Trichoplusia* larvae. Sequences encoding PTPase-like protein polypeptides can be cloned into a non-essential region of the virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of PTPase-like protein polypeptides will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein. The recombinant viruses can then be used to infect *S. frugiperda cells* or *Trichoplusia* larvae in which PTPase-like protein polypeptides can be expressed (Engelhard *et al.*, *Proc. Nat. Acad. Sci. 91*, 3224-3227, 1994).

Mammalian Expression Systems

A number of viral-based expression systems can be used to express PTPase-like protein polypeptides in mammalian host cells. For example, if an adenovirus is used as an expression vector, sequences encoding PTPase-like protein polypeptides can be ligated into an adenovirus transcription/translation complex comprising the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome can be used to obtain a viable virus that is capable of expressing a PTPase-like protein polypeptide in infected host cells (Logan & Shenk, *Proc. Natl.*

Acad. Sci. 81, 3655-3659, 1984). If desired, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, can be used to increase expression in mammalian host cells.

Human artificial chromosomes (HACs) also can be used to deliver larger fragments of DNA than can be contained and expressed in a plasmid. HACs of 6M to 10M are constructed and delivered to cells via conventional delivery methods (*e.g.*, liposomes, polycationic amino polymers, or vesicles).

Specific initiation signals also can be used to achieve more efficient translation of sequences encoding PTPase-like protein polypeptides. Such signals include the ATG initiation codon and adjacent sequences. In cases where sequences encoding a PTPase-like protein polypeptide, its initiation codon, and upstream sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals (including the ATG initiation codon) should be provided. The initiation codon should be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons can be of various origins, both natural and synthetic. The efficiency of expression can be enhanced by the inclusion of enhancers which are appropriate for the particular cell system which is used (see Scharf *et al.*, *Results Probl. Cell Differ. 20*, 125-162, 1994).

Host Cells

A host cell strain can be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed PTPase-like protein polypeptide in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the polypeptide also can be used to facilitate correct insertion, folding and/or function.

Different host cells that have specific cellular machinery and characteristic mechanisms for post-translational activities (*e.g.*, CHO, HeLa, MDCK, HEK293, and WI38), are available from the American Type Culture Collection (ATCC; 10801 University Boulevard, Manassas, VA 20110-2209) and can be chosen to ensure the correct modification and processing of the foreign protein.

Stable expression is preferred for long-term, high-yield production of recombinant proteins. For example, cell lines which stably express PTPase-like protein polypeptides can be transformed using expression vectors which can contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells can be allowed to grow for 1-2 days in an enriched medium before they are switched to a selective medium. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced PTPase-like protein sequences. Resistant clones of stably transformed cells can be proliferated using tissue culture techniques appropriate to the cell type. See, for example, ANIMAL CELL CULTURE, R.I. Freshney, ed., 1986.

Any number of selection systems can be used to recover transformed cell lines.

These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler *et al.*, *Cell 11*, 223-32, 1977) and adenine phosphoribosyltransferase (Lowy *et al.*, *Cell 22*, 817-23, 1980) genes which can be employed in *tk*$^-$ or *aprt*$^-$ cells, respectively. Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate (Wigler *et al.*, *Proc. Natl. Acad. Sci. 77*, 3567-70, 1980), *npt* confers resistance to the aminoglycosides, neomycin and G-418 (Colbere-Garapin *et al.*, *J. Mol. Biol. 150*, 1-14, 1981), and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Murray, 1992, *supra*). Additional selectable genes have been described. For example, *trpB* allows cells to utilize indole in place of

tryptophan, or *hisD*, which allows cells to utilize histinol in place of histidine (Hartman & Mulligan, *Proc. Natl. Acad. Sci. 85*, 8047-51, 1988). Visible markers such as anthocyanins, β-glucuronidase and its substrate GUS, and luciferase and its substrate luciferin, can be used to identify transformants and to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes *et al., Methods Mol. Biol. 55*, 121-131, 1995).

Detecting Expression

Although the presence of marker gene expression suggests that the PTPase-like protein polynucleotide is also present, its presence and expression may need to be confirmed. For example, if a sequence encoding a PTPase-like protein polypeptide is inserted within a marker gene sequence, transformed cells containing sequences that encode a PTPase-like protein polypeptide can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding a PTPase-like protein polypeptide under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the PTPase-like protein polynucleotide.

Alternatively, host cells which contain a PTPase-like protein polynucleotide and which express a PTPase-like protein polypeptide can be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations and protein bioassay or immunoassay techniques that include membrane, solution, or chip-based technologies for the detection and/or quantification of nucleic acid or protein. For example, the presence of a polynucleotide sequence encoding a PTPase-like protein polypeptide can be detected by DNA-DNA or DNA-RNA hybridization or amplification using probes or fragments or fragments of polynucleotides encoding a PTPase-like protein polypeptide. Nucleic acid amplification-based assays involve the use of oligonucleotides selected from sequences encoding a PTPase-like protein

polypeptide to detect transformants that contain a PTPase-like protein polynucleotide.

A variety of protocols for detecting and measuring the expression of a PTPase-like protein polypeptide, using either polyclonal or monoclonal antibodies specific for the polypeptide, are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay using monoclonal antibodies reactive to two non-interfering epitopes on a PTPase-like protein polypeptide can be used, or a competitive binding assay can be employed. These and other assays are described in Hampton *et al.*, SEROLOGICAL METHODS: A LABORATORY MANUAL, APS Press, St. Paul, Minn., 1990) and Maddox *et al.*, *J. Exp. Med. 158*, 1211-1216, 1983).

A wide variety of labels and conjugation techniques are known by those skilled in the art and can be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding PTPase-like protein polypeptides include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, sequences encoding a PTPase-like protein polypeptide can be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and can be used to synthesize RNA probes *in vitro* by addition of labeled nucleotides and an appropriate RNA polymerase such as T7, T3, or SP6. These procedures can be conducted using a variety of commercially available kits (Amersham Pharmacia Biotech, Promega, and US Biochemical). Suitable reporter molecules or labels which can be used for ease of detection include radionuclides, enzymes, and fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Expression and Purification of Polypeptides

Host cells transformed with nucleotide sequences encoding a PTPase-like protein polypeptide can be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The polypeptide produced by a transformed cell can be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode PTPase-like protein polypeptides can be designed to contain signal sequences which direct secretion of soluble PTPase-like protein polypeptides through a prokaryotic or eukaryotic cell membrane or which direct the membrane insertion of membrane-bound PTPase-like protein polypeptide.

As discussed above, other constructions can be used to join a sequence encoding a PTPase-like protein polypeptide to a nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, Wash.). Inclusion of cleavable linker sequences such as those specific for Factor Xa or enterokinase (Invitrogen, San Diego, CA) between the purification domain and the PTPase-like protein polypeptide also can be used to facilitate purification. One such expression vector provides for expression of a fusion protein containing a PTPase-like protein polypeptide and 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification by IMAC (immobilized metal ion affinity chromatography, as described in Porath *et al., Prot. Exp. Purif. 3*, 263-281, 1992), while the enterokinase cleavage site provides a means for purifying the PTPase-like protein polypeptide from the fusion protein. Vectors that contain fusion proteins are disclosed in Kroll *et al., DNA Cell Biol. 12*, 441-453, 1993.

Chemical Synthesis

Sequences encoding a PTPase-like protein polypeptide can be synthesized, in whole or in part, using chemical methods well known in the art (see Caruthers *et al.*, *Nucl. Acids Res. Symp. Ser.* 215-223, 1980; Horn *et al. Nucl. Acids Res. Symp. Ser.* 225-232, 1980). Alternatively, a PTPase-like protein polypeptide itself can be produced using chemical methods to synthesize its amino acid sequence, such as by direct peptide synthesis using solid-phase techniques (Merrifield, *J. Am. Chem. Soc. 85*, 2149-2154, 1963; Roberge *et al.*, *Science 269*, 202-204, 1995). Protein synthesis can be performed using manual techniques or by automation. Automated synthesis can be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer). Optionally, fragments of PTPase-like protein polypeptides can be separately synthesized and combined using chemical methods to produce a full-length molecule.

The newly synthesized peptide can be substantially purified by preparative high performance liquid chromatography (*e.g.*, Creighton, PROTEINS: STRUCTURES AND MOLECULAR PRINCIPLES, WH Freeman and Co., New York, N.Y., 1983). The composition of a synthetic PTPase-like protein polypeptide can be confirmed by amino acid analysis or sequencing (*e.g.*, the Edman degradation procedure; *see* Creighton, *supra*). Additionally, any portion of the amino acid sequence of the PTPase-like protein polypeptide can be altered during direct synthesis and/or combined using chemical methods with sequences from other proteins to produce a variant polypeptide or a fusion protein.

Production of Altered Polypeptides

As will be understood by those of skill in the art, it may be advantageous to produce PTPase-like protein polypeptide-encoding nucleotide sequences possessing non-naturally occurring codons. For example, codons preferred by a particular

prokaryotic or eukaryotic host can be selected to increase the rate of protein expression or to produce an RNA transcript having desirable properties, such as a half-life that is longer than that of a transcript generated from the naturally occurring sequence.

The nucleotide sequences disclosed herein can be engineered using methods generally known in the art to alter PTPase-like protein polypeptide-encoding sequences for a variety of reasons, including but not limited to, alterations which modify the cloning, processing, and/or expression of the polypeptide or mRNA product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides can be used to engineer the nucleotide sequences. For example, site-directed mutagenesis can be used to insert new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, introduce mutations, and so forth.

Antibodies

Any type of antibody known in the art can be generated to bind specifically to an epitope of a PTPase-like protein polypeptide. "Antibody" as used herein includes intact immunoglobulin molecules, as well as fragments thereof, such as Fab, $F(ab')_2$, and Fv, which are capable of binding an epitope of a PTPase-like protein polypeptide. Typically, at least 6, 8, 10, or 12 contiguous amino acids are required to form an epitope. However, epitopes which involve non-contiguous amino acids may require more, *e.g.*, at least 15, 25, or 50 amino acids.

An antibody which specifically binds to an epitope of a PTPase-like protein polypeptide can be used therapeutically, as well as in immunochemical assays, such as Western blots, ELISAs, radioimmunoassays, immunohistochemical assays, immunoprecipitations, or other immunochemical assays known in the art. Various immunoassays can be used to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays are well

known in the art. Such immunoassays typically involve the measurement of complex formation between an immunogen and an antibody that specifically binds to the immunogen.

Typically, an antibody which specifically binds to a PTPase-like protein polypeptide provides a detection signal at least 5-, 10-, or 20-fold higher than a detection signal provided with other proteins when used in an immunochemical assay. Preferably, antibodies which specifically bind to PTPase-like protein polypeptides do not detect other proteins in immunochemical assays and can immunoprecipitate a PTPase-like protein polypeptide from solution.

Human PTPase-like protein polypeptides can be used to immunize a mammal, such as a mouse, rat, rabbit, guinea pig, monkey, or human, to produce polyclonal antibodies. If desired, a PTPase-like protein polypeptide can be conjugated to a carrier protein, such as bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin. Depending on the host species, various adjuvants can be used to increase the immunological response. Such adjuvants include, but are not limited to, Freund's adjuvant, mineral gels (*e.g.*, aluminum hydroxide), and surface active substances (*e.g.* lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, and dinitrophenol). Among adjuvants used in humans, BCG (*bacilli Calmette-Guerin*) and *Corynebacterium parvum* are especially useful.

Monoclonal antibodies that specifically bind to a PTPase-like protein polypeptide can be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These techniques include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique (Kohler *et al., Nature 256*, 495-497, 1985; Kozbor *et al., J. Immunol. Methods 81*, 31-42, 1985; Cote *et al., Proc. Natl. Acad. Sci. 80*, 2026-2030, 1983; Cole *et al., Mol. Cell Biol. 62*, 109-120, 1984).

In addition, techniques developed for the production of "chimeric antibodies," the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used (Morrison *et al.*, *Proc. Natl. Acad. Sci. 81*, 6851-6855, 1984; Neuberger *et al.*, *Nature 312*, 604-608, 1984; Takeda *et al.*, *Nature 314*, 452-454, 1985). Monoclonal and other antibodies also can be "humanized" to prevent a patient from mounting an immune response against the antibody when it is used therapeutically. Such antibodies may be sufficiently similar in sequence to human antibodies to be used directly in therapy or may require alteration of a few key residues. Sequence differences between rodent antibodies and human sequences can be minimized by replacing residues which differ from those in the human sequences by site directed mutagenesis of individual residues or by grating of entire complementarity determining regions. Alternatively, humanized antibodies can be produced using recombinant methods, as described in GB2188638B. Antibodies that specifically bind to a PTPase-like protein polypeptide can contain antigen binding sites which are either partially or fully humanized, as disclosed in U.S. 5,565,332.

Alternatively, techniques described for the production of single chain antibodies can be adapted using methods known in the art to produce single chain antibodies that specifically bind to PTPase-like protein polypeptides. Antibodies with related specificity, but of distinct idiotypic composition, can be generated by chain shuffling from random combinatorial immunoglobin libraries (Burton, *Proc. Natl. Acad. Sci. 88*, 11120-23, 1991).

Single-chain antibodies also can be constructed using a DNA amplification method, such as PCR, using hybridoma cDNA as a template (Thirion *et al.*, 1996, *Eur. J. Cancer Prev. 5*, 507-11). Single-chain antibodies can be mono- or bispecific, and can be bivalent or tetravalent. Construction of tetravalent, bispecific single-chain antibodies is taught, for example, in Coloma & Morrison, 1997, *Nat. Biotechnol. 15*, 159-63. Construction of bivalent, bispecific single-chain antibodies is taught in Mallender & Voss, 1994, *J. Biol. Chem. 269*, 199-206.

A nucleotide sequence encoding a single-chain antibody can be constructed using manual or automated nucleotide synthesis, cloned into an expression construct using standard recombinant DNA methods, and introduced into a cell to express the coding sequence, as described below. Alternatively, single-chain antibodies can be produced directly using, for example, filamentous phage technology (Verhaar *et al.*, 1995, *Int. J. Cancer 61*, 497-501; Nicholls *et al.*, 1993, *J. Immunol. Meth. 165*, 81-91).

Antibodies which specifically bind to PTPase-like protein polypeptides also can be produced by inducing *in vivo* production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature (Orlandi *et al.*, *Proc. Natl. Acad. Sci. 86*, 3833-3837, 1989; Winter *et al.*, *Nature 349*, 293-299, 1991).

Other types of antibodies can be constructed and used therapeutically in methods of the invention. For example, chimeric antibodies can be constructed as disclosed in WO 93/03151. Binding proteins which are derived from immunoglobulins and which are multivalent and multispecific, such as the "diabodies" described in WO 94/13804, also can be prepared.

Antibodies according to the invention can be purified by methods well known in the art. For example, antibodies can be affinity purified by passage over a column to which a PTPase-like protein polypeptide is bound. The bound antibodies can then be eluted from the column using a buffer with a high salt concentration.

Antisense Oligonucleotides

Antisense oligonucleotides are nucleotide sequences that are complementary to a specific DNA or RNA sequence. Once introduced into a cell, the complementary nucleotides combine with natural sequences produced by the cell to form complexes and block either transcription or translation. Preferably, an antisense oligonucleotide

is at least 11 nucleotides in length, but can be at least 12, 15, 20, 25, 30, 35, 40, 45, or 50 or more nucleotides long. Longer sequences also can be used. Antisense oligonucleotide molecules can be provided in a DNA construct and introduced into a cell as described above to decrease the level of PTPase-like protein gene products in the cell.

Antisense oligonucleotides can be deoxyribonucleotides, ribonucleotides, or a combination of both. Oligonucleotides can be synthesized manually or by an automated synthesizer, by covalently linking the 5' end of one nucleotide with the 3' end of another nucleotide with non-phosphodiester internucleotide linkages such alkylphosphonates, phosphorothioates, phosphorodithioates, alkylphosphonothioates, alkylphosphonates, phosphoramidates, phosphate esters, carbamates, acetamidate, carboxymethyl esters, carbonates, and phosphate triesters. *See* Brown, *Meth. Mol. Biol. 20*, 1-8, 1994; Sonveaux, *Meth. Mol. Biol. 26*, 1-72, 1994; Uhlmann *et al.*, *Chem. Rev. 90*, 543-583, 1990.

Modifications of PTPase-like protein gene expression can be obtained by designing antisense oligonucleotides that will form duplexes to the control, 5', or regulatory regions of the PTPase-like protein gene. Oligonucleotides derived from the transcription initiation site, *e.g.*, between positions -10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using "triple helix" base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or chaperons. Therapeutic advances using triplex DNA have been described in the literature (*e.g.*, Gee *et al.*, in Huber & Carr, MOLECULAR AND IMMUNOLOGIC APPROACHES, Futura Publishing Co., Mt. Kisco, N.Y., 1994). An antisense oligonucleotide also can be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Precise complementarity is not required for successful complex formation between an antisense oligonucleotide and the complementary sequence of a PTPase-like

protein polynucleotide. Antisense oligonucleotides which comprise, for example, 2, 3, 4, or 5 or more stretches of contiguous nucleotides which are precisely complementary to a PTPase-like protein polynucleotide, each separated by a stretch of contiguous nucleotides which are not complementary to adjacent PTPase-like protein nucleotides, can provide sufficient targeting specificity for PTPase-like protein mRNA. Preferably, each stretch of complementary contiguous nucleotides is at least 4, 5, 6, 7, or 8 or more nucleotides in length. Non-complementary intervening sequences are preferably 1, 2, 3, or 4 nucleotides in length. One skilled in the art can easily use the calculated melting point of an antisense-sense pair to determine the degree of mismatching which will be tolerated between a particular antisense oligonucleotide and a particular PTPase-like protein polynucleotide sequence.

Antisense oligonucleotides can be modified without affecting their ability to hybridize to a PTPase-like protein polynucleotide. These modifications can be internal or at one or both ends of the antisense molecule. For example, internucleoside phosphate linkages can be modified by adding cholesteryl or diamine moieties with varying numbers of carbon residues between the amino groups and terminal ribose. Modified bases and/or sugars, such as arabinose instead of ribose, or a 3', 5'-substituted oligonucleotide in which the 3' hydroxyl group or the 5' phosphate group are substituted, also can be employed in a modified antisense oligonucleotide. These modified oligonucleotides can be prepared by methods well known in the art. *See, e.g.*, Agrawal *et al.*, *Trends Biotechnol. 10*, 152-158, 1992; Uhlmann *et al.*, *Chem. Rev. 90*, 543-584, 1990; Uhlmann *et al.*, *Tetrahedron. Lett. 215*, 3539-3542, 1987.

Ribozymes

Ribozymes are RNA molecules with catalytic activity. *See, e.g.*, Cech, *Science 236*, 1532-1539; 1987; Cech, *Ann. Rev. Biochem. 59*, 543-568; 1990, Cech, *Curr. Opin. Struct. Biol. 2*, 605-609; 1992, Couture & Stinchcomb, *Trends Genet. 12*, 510-515,

1996. Ribozymes can be used to inhibit gene function by cleaving an RNA sequence, as is known in the art (*e.g.*, Haseloff *et al.*, U.S. Patent 5,641,673). The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. Examples include engineered hammerhead motif ribozyme molecules that can specifically and efficiently catalyze endonucleolytic cleavage of specific nucleotide sequences.

The coding sequence of a PTPase-like protein polynucleotide can be used to generate ribozymes that will specifically bind to mRNA transcribed from the PTPase-like protein polynucleotide. Methods of designing and constructing ribozymes which can cleave other RNA molecules in trans in a highly sequence specific manner have been developed and described in the art (*see* Haseloff *et al. Nature 334*, 585-591, 1988). For example, the cleavage activity of ribozymes can be targeted to specific RNAs by engineering a discrete "hybridization" region into the ribozyme. The hybridization region contains a sequence complementary to the target RNA and thus specifically hybridizes with the target (see, for example, Gerlach *et al.*, EP 321,201).

Specific ribozyme cleavage sites within a PTPase-like protein RNA target can be identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the target RNA containing the cleavage site can be evaluated for secondary structural features which may render the target inoperable. Suitability of candidate PTPase-like protein RNA targets also can be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays. Longer complementary sequences can be used to increase the affinity of the hybridization sequence for the target. The hybridizing and cleavage regions of the ribozyme can be integrally related such that upon hybridizing to the target RNA through the complementary regions, the catalytic region of the ribozyme can cleave the target.

Ribozymes can be introduced into cells as part of a DNA construct. Mechanical methods, such as microinjection, liposome-mediated transfection, electroporation, or calcium phosphate precipitation, can be used to introduce a ribozyme-containing DNA construct into cells in which it is desired to decrease PTPase-like protein expression. Alternatively, if it is desired that the cells stably retain the DNA construct, the construct can be supplied on a plasmid and maintained as a separate element or integrated into the genome of the cells, as is known in the art. A ribozyme-encoding DNA construct can include transcriptional regulatory elements, such as a promoter element, an enhancer or UAS element, and a transcriptional terminator signal, for controlling transcription of ribozymes in the cells.

As taught in Haseloff *et al.*, U.S. Patent 5,641,673, ribozymes can be engineered so that ribozyme expression will occur in response to factors that induce expression of a target gene. Ribozymes also can be engineered to provide an additional level of regulation, so that destruction of mRNA occurs only when both a ribozyme and a target gene are induced in the cells.

Differentially Expressed Genes

Described herein are methods for the identification of genes whose products interact with human PTPase-like protein. Such genes may represent genes that are differentially expressed in disorders including, but not limited to, obesity, diabetes, CNS disorders, cardiovascular disorders, and COPD.

Further, such genes may represent genes that are differentially regulated in response to manipulations relevant to the progression or treatment of such diseases. Additionally, such genes may have a temporally modulated expression, increased or decreased at different stages of tissue or organism development. A differentially expressed gene may also have its expression modulated under control versus experimental conditions. In addition, the human PTPase-like protein gene or gene product may itself be tested for differential expression.

The degree to which expression differs in a normal versus a diseased state need only be large enough to be visualized via standard characterization techniques such as differential display techniques. Other such standard characterization techniques by which expression differences may be visualized include but are not limited to, quantitative RT (reverse transcriptase), PCR, and Northern analysis.

Identification of Differentially Expressed Genes

To identify differentially expressed genes total RNA or, preferably, mRNA is isolated from tissues of interest. For example, RNA samples are obtained from tissues of experimental subjects and from corresponding tissues of control subjects. Any RNA isolation technique that does not select against the isolation of mRNA may be utilized for the purification of such RNA samples. See, for example, Ausubel *et al.*, ed., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, Inc. New York, 1987-1993. Large numbers of tissue samples may readily be processed using techniques well known to those of skill in the art, such as, for example, the single-step RNA isolation process of Chomczynski, U.S. Patent 4,843,155.

Transcripts within the collected RNA samples that represent RNA produced by differentially expressed genes are identified by methods well known to those of skill in the art. They include, for example, differential screening (Tedder *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 85*, 208-12, 1988), subtractive hybridization (Hedrick *et al.*, *Nature 308*, 149-53; Lee *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 88*, 2825, 1984), and, preferably, differential display (Liang & Pardee, *Science 257*, 967-71, 1992; U.S. Patent 5,262,311).

The differential expression information may itself suggest relevant methods for the treatment of disorders involving the human PTPase-like protein. For example, treatment may include a modulation of expression of the differentially expressed genes and/or the gene encoding the human PTPase-like protein. The differential

expression information may indicate whether the expression or activity of the differentially expressed gene or gene product or the human PTPase-like protein gene or gene product are up-regulated or down-regulated.

Screening Methods

The invention provides assays for screening test compounds that bind to or modulate the activity of a PTPase-like protein polypeptide or a PTPase-like protein polynucleotide. A test compound preferably binds to a PTPase-like protein polypeptide or polynucleotide. More preferably, a test compound decreases or increases tyrosine phosphatase activity by at least about 10, preferably about 50, more preferably about 75, 90, or 100% relative to the absence of the test compound.

Test Compounds

Test compounds can be pharmacologic agents already known in the art or can be compounds previously unknown to have any pharmacological activity. The compounds can be naturally occurring or designed in the laboratory. They can be isolated from microorganisms, animals, or plants, and can be produced recombinantly, or synthesized by chemical methods known in the art. If desired, test compounds can be obtained using any of the numerous combinatorial library methods known in the art, including but not limited to, biological libraries, spatially addressable parallel solid phase or solution phase libraries, synthetic library methods requiring deconvolution, the "one-bead one-compound" library method, and synthetic library methods using affinity chromatography selection. The biological library approach is limited to polypeptide libraries, while the other four approaches are applicable to polypeptide, non-peptide oligomer, or small molecule libraries of compounds. *See* Lam, *Anticancer Drug Des. 12*, 145, 1997.

Methods for the synthesis of molecular libraries are well known in the art (*see*, for example, DeWitt *et al., Proc. Natl. Acad. Sci. U.S.A. 90*, 6909, 1993; Erb *et al. Proc.*

Natl. Acad. Sci. U.S.A. 91, 11422, 1994; Zuckermann *et al., J. Med. Chem. 37*, 2678, 1994; Cho *et al., Science 261*, 1303, 1993; Carell *et al., Angew. Chem. Int. Ed. Engl. 33*, 2059, 1994; Carell *et al., Angew. Chem. Int. Ed. Engl. 33*, 2061; Gallop *et al., J. Med. Chem. 37*, 1233, 1994). Libraries of compounds can be presented in solution (*see, e.g.*, Houghten, *BioTechniques 13*, 412-421, 1992), or on beads (Lam, *Nature 354*, 82-84, 1991), chips (Fodor, *Nature 364*, 555-556, 1993), bacteria or spores (Ladner, U.S. Patent 5,223,409), plasmids (Cull *et al., Proc. Natl. Acad. Sci. U.S.A. 89*, 1865-1869, 1992), or phage (Scott & Smith, *Science 249*, 386-390, 1990; Devlin, *Science 249*, 404-406, 1990); Cwirla *et al., Proc. Natl. Acad. Sci. 97*, 6378-6382, 1990; Felici, *J. Mol. Biol. 222*, 301-310, 1991; and Ladner, U.S. Patent 5,223,409).

High Throughput Screening

Test compounds can be screened for the ability to bind to PTPase-like protein polypeptides or polynucleotides or to affect PTPase-like protein activity or PTPase-like protein gene expression using high throughput screening. Using high throughput screening, many discrete compounds can be tested in parallel so that large numbers of test compounds can be quickly screened. The most widely established techniques utilize 96-well microtiter plates. The wells of the microtiter plates typically require assay volumes that range from 50 to 500 µl. In addition to the plates, many instruments, materials, pipettors, robotics, plate washers, and plate readers are commercially available to fit the 96-well format.

Alternatively, "free format assays," or assays that have no physical barrier between samples, can be used. For example, an assay using pigment cells (melanocytes) in a simple homogeneous assay for combinatorial peptide libraries is described by Jayawickreme *et al., Proc. Natl. Acad. Sci. U.S.A. 19*, 1614-18 (1994). The cells are placed under agarose in petri dishes, then beads that carry combinatorial compounds are placed on the surface of the agarose. The combinatorial compounds are partially released the compounds from the beads. Active compounds can be visualized as

dark pigment areas because, as the compounds diffuse locally into the gel matrix, the active compounds cause the cells to change colors.

Another example of a free format assay is described by Chelsky, "Strategies for Screening Combinatorial Libraries: Novel and Traditional Approaches," reported at the First Annual Conference of The Society for Biomolecular Screening in Philadelphia, Pa. (Nov. 7-10, 1995). Chelsky placed a simple homogenous enzyme assay for carbonic anhydrase inside an agarose gel such that the enzyme in the gel would cause a color change throughout the gel. Thereafter, beads carrying combinatorial compounds via a photolinker were placed inside the gel and the compounds were partially released by UV-light. Compounds that inhibited the enzyme were observed as local zones of inhibition having less color change.

Yet another example is described by Salmon *et al.*, *Molecular Diversity 2*, 57-63 (1996). In this example, combinatorial libraries were screened for compounds that had cytotoxic effects on cancer cells growing in agar.

Another high throughput screening method is described in Beutel *et al.*, U.S. Patent 5,976,813. In this method, test samples are placed in a porous matrix. One or more assay components are then placed within, on top of, or at the bottom of a matrix such as a gel, a plastic sheet, a filter, or other form of easily manipulated solid support. When samples are introduced to the porous matrix they diffuse sufficiently slowly, such that the assays can be performed without the test samples running together.

Binding Assays

For binding assays, the test compound is preferably a small molecule that binds to and occupies, for example, the active site of the PTPase-like protein polypeptide, such that normal biological activity is prevented. Examples of such small molecules include, but are not limited to, small peptides or peptide-like molecules.

In binding assays, either the test compound or the PTPase-like protein polypeptide can comprise a detectable label, such as a fluorescent, radioisotopic, chemiluminescent, or enzymatic label, such as horseradish peroxidase, alkaline phosphatase, or luciferase. Detection of a test compound that is bound to the PTPase-like protein polypeptide can then be accomplished, for example, by direct counting of radioemmission, by scintillation counting, or by determining conversion of an appropriate substrate to a detectable product.

Alternatively, binding of a test compound to a PTPase-like protein polypeptide can be determined without labeling either of the interactants. For example, a microphysiometer can be used to detect binding of a test compound with a PTPase-like protein polypeptide. A microphysiometer (*e.g.*, Cytosensor™) is an analytical instrument that measures the rate at which a cell acidifies its environment using a light-addressable potentiometric sensor (LAPS). Changes in this acidification rate can be used as an indicator of the interaction between a test compound and a PTPase-like protein polypeptide (McConnell *et al.*, *Science 257*, 1906-1912, 1992).

Determining the ability of a test compound to bind to a PTPase-like protein polypeptide also can be accomplished using a technology such as real-time Bimolecular Interaction Analysis (BIA) (Sjolander & Urbaniczky, *Anal. Chem. 63*, 2338-2345, 1991, and Szabo *et al.*, *Curr. Opin. Struct. Biol. 5*, 699-705, 1995). BIA is a technology for studying biospecific interactions in real time, without labeling any of the interactants (*e.g.*, BIAcore™). Changes in the optical phenomenon surface plasmon resonance (SPR) can be used as an indication of real-time reactions between biological molecules.

In yet another aspect of the invention, a PTPase-like protein polypeptide can be used as a "bait protein" in a two-hybrid assay or three-hybrid assay (see, *e.g.*, U.S. Patent 5,283,317; Zervos *et al.*, *Cell 72*, 223-232, 1993; Madura *et al.*, *J. Biol. Chem. 268*, 12046-12054, 1993; Bartel *et al.*, *BioTechniques 14*, 920-924, 1993; Iwabuchi *et al.*, *Oncogene 8*, 1693-1696, 1993; and Brent W094/10300), to identify other proteins

which bind to or interact with the PTPase-like protein polypeptide and modulate its activity.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. For example, in one construct, polynucleotide encoding a PTPase-like protein polypeptide can be fused to a polynucleotide encoding the DNA binding domain of a known transcription factor (*e.g.*, GAL-4). In the other construct a DNA sequence that encodes an unidentified protein ("prey" or "sample") can be fused to a polynucleotide that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact *in vivo* to form an protein-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.*, LacZ), which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected, and cell colonies containing the functional transcription factor can be isolated and used to obtain the DNA sequence encoding the protein that interacts with the PTPase-like protein polypeptide.

It may be desirable to immobilize either the PTPase-like protein polypeptide (or polynucleotide) or the test compound to facilitate separation of bound from unbound forms of one or both of the interactants, as well as to accommodate automation of the assay. Thus, either the PTPase-like protein polypeptide (or polynucleotide) or the test compound can be bound to a solid support. Suitable solid supports include, but are not limited to, glass or plastic slides, tissue culture plates, microtiter wells, tubes, silicon chips, or particles such as beads (including, but not limited to, latex, polystyrene, or glass beads). Any method known in the art can be used to attach the enzyme polypeptide (or polynucleotide) or test compound to a solid support, including use of covalent and non-covalent linkages, passive absorption, or pairs of binding moieties attached respectively to the polypeptide (or polynucleotide) or test compound and the solid support. Test compounds are preferably bound to the solid

support in an array, so that the location of individual test compounds can be tracked. Binding of a test compound to a PTPase-like protein polypeptide (or polynucleotide) can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtiter plates, test tubes, and microcentrifuge tubes.

In one embodiment, the PTPase-like protein polypeptide is a fusion protein comprising a domain that allows the PTPase-like protein polypeptide to be bound to a solid support. For example, glutathione-S-transferase fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, Mo.) or glutathione derivatized microtiter plates, which are then combined with the test compound or the test compound and the non-adsorbed PTPase-like protein polypeptide; the mixture is then incubated under conditions conducive to complex formation (*e.g.*, at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to remove any unbound components. Binding of the interactants can be determined either directly or indirectly, as described above. Alternatively, the complexes can be dissociated from the solid support before binding is determined.

Other techniques for immobilizing proteins or polynucleotides on a solid support also can be used in the screening assays of the invention. For example, either a PTPase-like protein polypeptide (or polynucleotide) or a test compound can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated PTPase-like protein polypeptides (or polynucleotides) or test compounds can be prepared from biotin-NHS(N-hydroxysuccinimide) using techniques well known in the art (*e.g.*, biotinylation kit, Pierce Chemicals, Rockford, Ill.) and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies which specifically bind to a PTPase-like protein polypeptide, polynucleotide, or a test compound, but which do not interfere with a desired binding site, such as the active site of the PTPase-like protein polypeptide, can be derivatized to the wells of the plate. Unbound target or protein can be trapped in the wells by antibody conjugation.

Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies which specifically bind to the PTPase-like protein polypeptide or test compound, enzyme-linked assays which rely on detecting an activity of the PTPase-like protein polypeptide, and SDS gel electrophoresis under non-reducing conditions.

Screening for test compounds which bind to a PTPase-like protein polypeptide or polynucleotide also can be carried out in an intact cell. Any cell which comprises a PTPase-like protein polypeptide or polynucleotide can be used in a cell-based assay system. A PTPase-like protein polynucleotide can be naturally occurring in the cell or can be introduced using techniques such as those described above. Binding of the test compound to a PTPase-like protein polypeptide or polynucleotide is determined as described above.

Enzyme Assays

Test compounds can be tested for the ability to increase or decrease the tyrosine phosphatase activity of a human PTPase-like protein polypeptide. Tyrosine phosphatase activity can be measured, for example, as described in Calvert-Evers & Hammond, *Cell Biol. Int. 24*, 559-68, 2000.

Enzyme assays can be carried out after contacting either a purified PTPase-like protein polypeptide, a cell membrane preparation, or an intact cell with a test compound. A test compound that decreases a tyrosine phosphatase activity of a PTPase-like protein polypeptide by at least about 10, preferably about 50, more preferably about 75, 90, or 100% is identified as a potential therapeutic agent for decreasing PTPase-like protein activity. A test compound which increases a tyrosine phosphatase activity of a human PTPase-like protein polypeptide by at least about 10, preferably about 50, more preferably about 75, 90, or 100% is identified as a potential therapeutic agent for increasing human PTPase-like protein activity.

Gene Expression

In another embodiment, test compounds that increase or decrease PTPase-like protein gene expression are identified. A PTPase-like protein polynucleotide is contacted with a test compound, and the expression of an RNA or polypeptide product of the PTPase-like protein polynucleotide is determined. The level of expression of appropriate mRNA or polypeptide in the presence of the test compound is compared to the level of expression of mRNA or polypeptide in the absence of the test compound. The test compound can then be identified as a modulator of expression based on this comparison. For example, when expression of mRNA or polypeptide is greater in the presence of the test compound than in its absence, the test compound is identified as a stimulator or enhancer of the mRNA or polypeptide expression. Alternatively, when expression of the mRNA or polypeptide is less in the presence of the test compound than in its absence, the test compound is identified as an inhibitor of the mRNA or polypeptide expression.

The level of PTPase-like protein mRNA or polypeptide expression in the cells can be determined by methods well known in the art for detecting mRNA or polypeptide. Either qualitative or quantitative methods can be used. The presence of polypeptide products of a PTPase-like protein polynucleotide can be determined, for example, using a variety of techniques known in the art, including immunochemical methods such as radioimmunoassay, Western blotting, and immunohistochemistry. Alternatively, polypeptide synthesis can be determined *in vivo*, in a cell culture, or in an *in vitro* translation system by detecting incorporation of labeled amino acids into a PTPase-like protein polypeptide.

Such screening can be carried out either in a cell-free assay system or in an intact cell. Any cell that expresses a PTPase-like protein polynucleotide can be used in a cell-based assay system. The PTPase-like protein polynucleotide can be naturally occurring in the cell or can be introduced using techniques such as those described

above. Either a primary culture or an established cell line, such as CHO or human embryonic kidney 293 cells, can be used.

Pharmaceutical Compositions

The invention also provides pharmaceutical compositions that can be administered to a patient to achieve a therapeutic effect. Pharmaceutical compositions of the invention can comprise, for example, a PTPase-like protein polypeptide, PTPase-like protein polynucleotide, ribozymes or antisense oligonucleotides, antibodies which specifically bind to a PTPase-like protein polypeptide, or mimetics, activators, or inhibitors of a PTPase-like protein polypeptide activity. The compositions can be administered alone or in combination with at least one other agent, such as stabilizing compound, which can be administered in any sterile, biocompatible pharmaceutical carrier, including, but not limited to, saline, buffered saline, dextrose, and water. The compositions can be administered to a patient alone, or in combination with other agents, drugs or hormones.

In addition to the active ingredients, these pharmaceutical compositions can contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries that facilitate processing of the active compounds into preparations which can be used pharmaceutically. Pharmaceutical compositions of the invention can be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, parenteral, topical, sublingual, or rectal means. Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combination of active compounds with solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, or sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums including arabic and tragacanth; and proteins such as gelatin and collagen. If desired, disintegrating or solubilizing agents can be added, such as the cross-linked polyvinyl pyrrolidone, agar, alginic acid, or a salt thereof, such as sodium alginate.

Dragee cores can be used in conjunction with suitable coatings, such as concentrated sugar solutions, which also can contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments can be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, *i.e.*, dosage.

Pharmaceutical preparations that can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with a filler or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds can be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration can be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions can contain substances that increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions

of the active compounds can be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Non-lipid polycationic amino polymers also can be used for delivery. Optionally, the suspension also can contain suitable stabilizers or agents that increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention can be manufactured in a manner that is known in the art, *e.g.*, by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes. The pharmaceutical composition can be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, succinic, etc. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation can be a lyophilized powder which can contain any or all of the following: 1-50 mM histidine, 0.1%-2% sucrose, and 2-7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

Further details on techniques for formulation and administration can be found in the latest edition of REMINGTON'S PHARMACEUTICAL SCIENCES (Maack Publishing Co., Easton, Pa.). After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. Such labeling would include amount, frequency, and method of administration.

Therapeutic Indications and Methods

Human PTPase-like protein can be regulated to treat, obesity, diabetes, CNS disorders, cardiovascular disorders, and COPD.

Obesity and overweight are defined as an excess of body fat relative to lean body mass. An increase in caloric intake or a decrease in energy expenditure or both can bring about this imbalance leading to surplus energy being stored as fat. Obesity is associated with important medical morbidities and an increase in mortality. The causes of obesity are poorly understood and may be due to genetic factors, environmental factors or a combination of the two to cause a positive energy balance. In contrast, anorexia and cachexia are characterized by an imbalance in energy intake versus energy expenditure leading to a negative energy balance and weight loss. Agents that either increase energy expenditure and/or decrease energy intake, absorption or storage would be useful for treating obesity, overweight, and associated comorbidities. Agents that either increase energy intake and/or decrease energy expenditure or increase the amount of lean tissue would be useful for treating cachexia, anorexia and wasting disorders.

This gene, translated proteins and agents which modulate this gene or portions of the gene or its products are useful for treating obesity, overweight, anorexia, cachexia, wasting disorders, appetite suppression, appetite enhancement, increases or decreases in satiety, modulation of body weight, and/or other eating disorders such as bulimia. Also this gene, translated proteins and agents which modulate this gene or portions of the gene or its products are useful for treating obesity/overweight-associated comorbidities including hypertension, type 2 diabetes, coronary artery disease, hyperlipidemia, stroke, gallbladder disease, gout, osteoarthritis, sleep apnea and respiratory problems, some types of cancer including endometrial, breast, prostate, and colon cancer, thrombolic disease, polycystic ovarian syndrome, reduced fertility, complications of pregnancy, menstrual irregularities, hirsutism, stress incontinence, and depression.

Diabetes mellitus is a common metabolic disorder characterized by an abnormal elevation in blood glucose, alterations in lipids and abnormalities (complications) in the cardiovascular system, eye, kidney and nervous system. Diabetes is divided into two separate diseases: type 1 diabetes (juvenile onset), which results from a loss of cells which make and secrete insulin, and type 2 diabetes (adult onset), which is caused by a defect in insulin secretion and a defect in insulin action.

Type 1 diabetes is initiated by an autoimmune reaction that attacks the insulin secreting cells (beta cells) in the pancreatic islets. Agents that prevent this reaction from occurring or that stop the reaction before destruction of the beta cells has been accomplished are potential therapies for this disease. Other agents that induce beta cell proliferation and regeneration also are potential therapies.

Type II diabetes is the most common of the two diabetic conditions (6% of the population). The defect in insulin secretion is an important cause of the diabetic condition and results from an inability of the beta cell to properly detect and respond to rises in blood glucose levels with insulin release. Therapies that increase the response by the beta cell to glucose would offer an important new treatment for this disease.

The defect in insulin action in Type II diabetic subjects is another target for therapeutic intervention. Agents that increase the activity of the insulin receptor in muscle, liver, and fat will cause a decrease in blood glucose and a normalization of plasma lipids. The receptor activity can be increased by agents that directly stimulate the receptor or that increase the intracellular signals from the receptor. Other therapies can directly activate the cellular end process, *i.e.* glucose transport or various enzyme systems, to generate an insulin-like effect and therefore a produce beneficial outcome. Because overweight subjects have a greater susceptibility to Type II diabetes, any agent that reduces body weight is a possible therapy.

Both Type I and Type diabetes can be treated with agents that mimic insulin action or that treat diabetic complications by reducing blood glucose levels. Likewise, agents that reduces new blood vessel growth can be used to treat the eye complications that develop in both diseases.

CNS disorders. Central and peripheral nervous system disorders also can be treated, such as primary and secondary disorders after brain injury, disorders of mood, anxiety disorders, disorders of thought and volition, disorders of sleep and wakefulness, diseases of the motor unit, such as neurogenic and myopathic disorders, neurodegenerative disorders such as Alzheimer's and Parkinson's disease, and processes of peripheral and chronic pain.

Pain that is associated with CNS disorders also can be treated by regulating the activity of human PTPase-like protein. Pain which can be treated includes that associated with central nervous system disorders, such as multiple sclerosis, spinal cord injury, sciatica, failed back surgery syndrome, traumatic brain injury, epilepsy, Parkinson's disease, post-stroke, and vascular lesions in the brain and spinal cord (*e.g.*, infarct, hemorrhage, vascular malformation). Non-central neuropathic pain includes that associated with post mastectomy pain, reflex sympathetic dystrophy (RSD), trigeminal neuralgiaradioculopathy, post-surgical pain, HIV/AIDS related pain, cancer pain, metabolic neuropathies (*e.g.*, diabetic neuropathy, vasculitic neuropathy secondary to connective tissue disease), paraneoplastic polyneuropathy associated, for example, with carcinoma of lung, or leukemia, or lymphoma, or carcinoma of prostate, colon or stomach, trigeminal neuralgia, cranial neuralgias, and post-herpetic neuralgia. Pain associated with cancer and cancer treatment also can be treated, as can headache pain (for example, migraine with aura, migraine without aura, and other migraine disorders), episodic and chronic tension-type headache, tension-type like headache, cluster headache, and chronic paroxysmal hemicrania.

This invention further pertains to the use of novel agents identified by the screening assays described above. Accordingly, it is within the scope of this invention to use a test compound identified as described herein in an appropriate animal model.

Cardiovascular disorders. Cardiovascular diseases include the following disorders of the heart and the vascular system: congestive heart failure, myocardial infarction, ischemic diseases of the heart, all kinds of atrial and ventricular arrhythmias, hypertensive vascular diseases, and peripheral vascular diseases.

Heart failure is defined as a pathophysiologic state in which an abnormality of cardiac function is responsible for the failure of the heart to pump blood at a rate commensurate with the requirement of the metabolizing tissue. It includes all forms of pumping failure, such as high-output and low-output, acute and chronic, right-sided or left-sided, systolic or diastolic, independent of the underlying cause.

Myocardial infarction (MI) is generally caused by an abrupt decrease in coronary blood flow that follows a thrombotic occlusion of a coronary artery previously narrowed by arteriosclerosis. MI prophylaxis (primary and secondary prevention) is included, as well as the acute treatment of MI and the prevention of complications.

Ischemic diseases are conditions in which the coronary flow is restricted resulting in a perfusion which is inadequate to meet the myocardial requirement for oxygen. This group of diseases includes stable angina, unstable angina, and asymptomatic ischemia.

Arrhythmias include all forms of atrial and ventricular tachyarrhythmias (atrial tachycardia, atrial flutter, atrial fibrillation, atrio-ventricular reentrant tachycardia, preexcitation syndrome, ventricular tachycardia, ventricular flutter, and ventricular fibrillation), as well as bradycardic forms of arrhythmias.

Vascular diseases include primary as well as all kinds of secondary arterial hypertension (renal, endocrine, neurogenic, others). The disclosed gene and its

product may be used as drug targets for the treatment of hypertension as well as for the prevention of all complications. Peripheral vascular diseases are defined as vascular diseases in which arterial and/or venous flow is reduced resulting in an imbalance between blood supply and tissue oxygen demand. It includes chronic peripheral arterial occlusive disease (PAOD), acute arterial thrombosis and embolism, inflammatory vascular disorders, Raynaud's phenomenon, and venous disorders.

<u>COPD.</u> Chronic obstructive pulmonary (or airways) disease (COPD) is a condition defined physiologically as airflow obstruction that generally results from a mixture of emphysema and peripheral airway obstruction due to chronic bronchitis (Senior & Shapiro, *Pulmonary Diseases and Disorders*, 3d ed., New York, McGraw-Hill, 1998, pp. 659-681, 1998; Barnes, *Chest 117*, 10S-14S, 2000). Emphysema is characterized by destruction of alveolar walls leading to abnormal enlargement of the air spaces of the lung. Chronic bronchitis is defined clinically as the presence of chronic productive cough for three months in each of two successive years. In COPD, airflow obstruction is usually progressive and is only partially reversible. By far the most important risk factor for development of COPD is cigarette smoking, although the disease does occur in non-smokers.

Chronic inflammation of the airways is a key pathological feature of COPD (Senior & Shapiro, 1998). The inflammatory cell population comprises increased numbers of macrophages, neutrophils, and $CD8^+$ lymphocytes. Inhaled irritants, such as cigarette smoke, activate macrophages which are resident in the respiratory tract, as well as epithelial cells leading to release of chemokines (*e.g.*, interleukin-8) and other chemotactic factors. These chemotactic factors act to increase the neutrophil/monocyte trafficking from the blood into the lung tissue and airways. Neutrophils and monocytes recruited into the airways can release a variety of potentially damaging mediators such as proteolytic enzymes and reactive oxygen species. Matrix degradation and emphysema, along with airway wall thickening, surfactant

dysfunction, and mucus hypersecretion, all are potential sequelae of this inflammatory response that lead to impaired airflow and gas exchange.

For example, an agent identified as described herein (*e.g.*, a modulating agent, an antisense nucleic acid molecule, a specific antibody, ribozyme, or a PTPase-like protein polypeptide binding molecule) can be used in an animal model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an agent identified as described herein can be used in an animal model to determine the mechanism of action of such an agent. Furthermore, this invention pertains to uses of novel agents identified by the above-described screening assays for treatments as described herein.

A reagent which affects PTPase-like protein activity can be administered to a human cell, either *in vitro* or *in vivo*, to reduce PTPase-like protein activity. The reagent preferably binds to an expression product of a human PTPase-like protein gene. If the expression product is a protein, the reagent is preferably an antibody. For treatment of human cells *ex vivo*, an antibody can be added to a preparation of stem cells that have been removed from the body. The cells can then be replaced in the same or another human body, with or without clonal propagation, as is known in the art.

In one embodiment, the reagent is delivered using a liposome. Preferably, the liposome is stable in the animal into which it has been administered for at least about 30 minutes, more preferably for at least about 1 hour, and even more preferably for at least about 24 hours. A liposome comprises a lipid composition that is capable of targeting a reagent, particularly a polynucleotide, to a particular site in an animal, such as a human. Preferably, the lipid composition of the liposome is capable of targeting to a specific organ of an animal, such as the lung, liver, spleen, heart brain, lymph nodes, and skin.

A liposome useful in the present invention comprises a lipid composition that is capable of fusing with the plasma membrane of the targeted cell to deliver its contents to the cell. Preferably, the transfection efficiency of a liposome is about 0.5 μg of DNA per 16 nmole of liposome delivered to about 10^6 cells, more preferably about 1.0 μg of DNA per 16 nmole of liposome delivered to about 10^6 cells, and even more preferably about 2.0 μg of DNA per 16 nmol of liposome delivered to about 10^6 cells. Preferably, a liposome is between about 100 and 500 nm, more preferably between about 150 and 450 nm, and even more preferably between about 200 and 400 nm in diameter.

Suitable liposomes for use in the present invention include those liposomes standardly used in, for example, gene delivery methods known to those of skill in the art. More preferred liposomes include liposomes having a polycationic lipid composition and/or liposomes having a cholesterol backbone conjugated to polyethylene glycol. Optionally, a liposome comprises a compound capable of targeting the liposome to a particular cell type, such as a cell-specific ligand exposed on the outer surface of the liposome.

Complexing a liposome with a reagent such as an antisense oligonucleotide or ribozyme can be achieved using methods that are standard in the art (see, for example, U.S. Patent 5,705,151). Preferably, from about 0.1 μg to about 10 μg of polynucleotide is combined with about 8 nmol of liposomes, more preferably from about 0.5 μg to about 5 μg of polynucleotides are combined with about 8 nmol liposomes, and even more preferably about 1.0 μg of polynucleotides is combined with about 8 nmol liposomes.

In another embodiment, antibodies can be delivered to specific tissues *in vivo* using receptor-mediated targeted delivery. Receptor-mediated DNA delivery techniques are taught in, for example, Findeis *et al. Trends in Biotechnol. 11*, 202-05 (1993); Chiou *et al.*, GENE THERAPEUTICS: METHODS AND APPLICATIONS OF DIRECT GENE TRANSFER (J.A. Wolff, ed.) (1994); Wu & Wu, *J. Biol. Chem. 263*, 621-24 (1988);

Wu *et al.*, *J. Biol. Chem. 269*, 542-46 (1994); Zenke *et al.*, *Proc. Natl. Acad. Sci. U.S.A. 87*, 3655-59 (1990); Wu *et al.*, *J. Biol. Chem. 266*, 338-42 (1991).

Determination of a Therapeutically Effective Dose

The determination of a therapeutically effective dose is well within the capability of those skilled in the art. A therapeutically effective dose refers to that amount of active ingredient which increases or decreases PTPase-like protein activity relative to the PTPase-like protein activity which occurs in the absence of the therapeutically effective dose.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays or in animal models, usually mice, rabbits, dogs, or pigs. The animal model also can be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

Therapeutic efficacy and toxicity, *e.g.*, ED_{50} (the dose therapeutically effective in 50% of the population) and LD_{50} (the dose lethal to 50% of the population), can be determined by standard pharmaceutical procedures in cell cultures or experimental animals. The dose ratio of toxic to therapeutic effects is the therapeutic index, and it can be expressed as the ratio, LD_{50}/ED_{50}.

Pharmaceutical compositions that exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies is used in formulating a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject that requires treatment. Dosage and administration are adjusted to provide sufficient levels of the active ingredient or to maintain the desired effect. Factors that can be taken into account include the severity of the disease state, general health of the subject, age, weight, and gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long-acting pharmaceutical compositions can be administered every 3 to 4 days, every week, or once every two weeks depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts can vary from 0.1 to 100,000 micrograms, up to a total dose of about 1 g, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

If the reagent is a single-chain antibody, polynucleotides encoding the antibody can be constructed and introduced into a cell either *ex vivo* or *in vivo* using well-established techniques including, but not limited to, transferrin-polycation-mediated DNA transfer, transfection with naked or encapsulated nucleic acids, liposome-mediated cellular fusion, intracellular transportation of DNA-coated latex beads, protoplast fusion, viral infection, electroporation, "gene gun," and DEAE- or calcium phosphate-mediated transfection.

Effective *in vivo* dosages of an antibody are in the range of about 5 μg to about 50 μg/kg, about 50 μg to about 5 mg/kg, about 100 μg to about 500 μg/kg of patient body weight, and about 200 to about 250 μg/kg of patient body weight. For administration of polynucleotides encoding single-chain antibodies, effective *in vivo* dosages are in the range of about 100 ng to about 200 ng, 500 ng to about 50 mg,

about 1 μg to about 2 mg, about 5 μg to about 500 μg, and about 20 μg to about 100 μg of DNA.

If the expression product is mRNA, the reagent is preferably an antisense oligonucleotide or a ribozyme. Polynucleotides that express antisense oligonucleotides or ribozymes can be introduced into cells by a variety of methods, as described above.

Preferably, a reagent reduces expression of a PTPase-like protein gene or the activity of a PTPase-like protein polypeptide by at least about 10, preferably about 50, more preferably about 75, 90, or 100% relative to the absence of the reagent. The effectiveness of the mechanism chosen to decrease the level of expression of a PTPase-like protein gene or the activity of a PTPase-like protein polypeptide can be assessed using methods well known in the art, such as hybridization of nucleotide probes to PTPase-like specific-specific mRNA, quantitative RT-PCR, immunologic detection of a PTPase-like protein polypeptide, or measurement of PTPase-like protein activity.

In any of the embodiments described above, any of the pharmaceutical compositions of the invention can be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy can be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents can act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

Any of the therapeutic methods described above can be applied to any subject in need of such therapy, including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

Diagnostic Methods

Human PTPase-like protein also can be used in diagnostic assays for detecting diseases and abnormalities or susceptibility to diseases and abnormalities related to the presence of mutations in the nucleic acid sequences that encode the enzyme. For example, differences can be determined between the cDNA or genomic sequence encoding PTPase-like protein in individuals afflicted with a disease and in normal individuals. If a mutation is observed in some or all of the afflicted individuals but not in normal individuals, then the mutation is likely to be the causative agent of the disease.

Sequence differences between a reference gene and a gene having mutations can be revealed by the direct DNA sequencing method. In addition, cloned DNA segments can be employed as probes to detect specific DNA segments. The sensitivity of this method is greatly enhanced when combined with PCR. For example, a sequencing primer can be used with a double-stranded PCR product or a single-stranded template molecule generated by a modified PCR. The sequence determination is performed by conventional procedures using radiolabeled nucleotides or by automatic sequencing procedures using fluorescent tags.

Genetic testing based on DNA sequence differences can be carried out by detection of alteration in electrophoretic mobility of DNA fragments in gels with or without denaturing agents. Small sequence deletions and insertions can be visualized, for example, by high resolution gel electrophoresis. DNA fragments of different sequences can be distinguished on denaturing formamide gradient gels in which the mobilities of different DNA fragments are retarded in the gel at different positions according to their specific melting or partial melting temperatures (*see, e.g.*, Myers *et al.*, *Science 230*, 1242, 1985). Sequence changes at specific locations can also be revealed by nuclease protection assays, such as RNase and S 1 protection or the chemical cleavage method (*e.g.*, Cotton *et al.*, *Proc. Natl. Acad. Sci. USA 85*, 4397-4401, 1985). Thus, the detection of a specific DNA sequence can be performed

by methods such as hybridization, RNase protection, chemical cleavage, direct DNA sequencing or the use of restriction enzymes and Southern blotting of genomic DNA. In addition to direct methods such as gel-electrophoresis and DNA sequencing, mutations can also be detected by *in situ* analysis.

Altered levels of PTPase-like protein also can be detected in various tissues. Assays used to detect levels of the receptor polypeptides in a body sample, such as blood or a tissue biopsy, derived from a host are well known to those of skill in the art and include radioimmunoassays, competitive binding assays, Western blot analysis, and ELISA assays.

All patents and patent applications cited in this disclosure are expressly incorporated herein by reference. The above disclosure generally describes the present invention. A more complete understanding can be obtained by reference to the following specific examples, which are provided for purposes of illustration only and are not intended to limit the scope of the invention.

EXAMPLE 1

Detection of PTPase-like protein activity

The polynucleotide of SEQ ID NO: 1 is inserted into the expression vector pCEV4 and the expression vector pCEV4-PTPase-like protein polypeptide obtained is transfected into human embryonic kidney 293 cells. From these cells extracts are obtained and tested in an assay comprising tyrosine-phosphorylated Raytide as substrate. Raytide is phosphorylated by incubation with c-Src and [gamma-^{32}P]ATP and is purified as previously described. Cell extracts are incubated with ^{32}P-Raytide in PTPase buffer (50 mM Hepes, pH 7.2, 60 mM NaCl, 60 mM KCl, and protease inhibitors) at a final volume of 60 µl. After incubation at 30°C for 10 min, the reaction is ended by spotting 50 µl of the reaction mixture to P81 phosphocellulose filter discs. Discs are washed in 0.5% phosphoric acid, dried, and placed in a scintillation counter to measure radioactivity. A decrease in filter-bound radioactivity is considered as PTPase activity. Zero minutes of incubation or presence of 1 mM vanadate is used for control assay. It is shown that the polypeptide of SEQ ID NO: 2 has a PTPase-like protein activity.

EXAMPLE 2

Expression of recombinant human PTPase-like protein

The *Pichia pastoris* expression vector pPICZB (Invitrogen, San Diego, CA) is used to produce large quantities of recombinant human PTPase-like protein polypeptides in yeast. The PTPase-like encoding-encoding DNA sequence is derived from SEQ ID NO: 1. Before insertion into vector pPICZB, the DNA sequence is modified by well known methods in such a way that it contains at its 5'-end an initiation codon and at its 3'-end an enterokinase cleavage site, a His6 reporter tag and a termination codon. Moreover, at both termini recognition sequences for restriction endonucleases are added and after digestion of the multiple cloning site of pPICZ B with the

corresponding restriction enzymes the modified DNA sequence is ligated into pPICZB. This expression vector is designed for inducible expression in *Pichia pastoris*, driven by a yeast promoter. The resulting pPICZ/md-His6 vector is used to transform the yeast.

The yeast is cultivated under usual conditions in 5 liter shake flasks and the recombinantly produced protein isolated from the culture by affinity chromatography (Ni-NTA-Resin) in the presence of 8 M urea. The bound polypeptide is eluted with buffer, pH 3.5, and neutralized. Separation of the polypeptide from the His6 reporter tag is accomplished by site-specific proteolysis using enterokinase (Invitrogen, San Diego, CA) according to manufacturer's instructions. Purified human PTPase-like protein polypeptide is obtained.

EXAMPLE 3

Identification of test compounds that bind to PTPase-like protein polypeptides

Purified PTPase-like protein polypeptides comprising a glutathione-S-transferase protein and absorbed onto glutathione-derivatized wells of 96-well microtiter plates are contacted with test compounds from a small molecule library at pH 7.0 in a physiological buffer solution. Human PTPase-like protein polypeptides comprise the amino acid sequence shown in SEQ ID NO: 2. The test compounds comprise a fluorescent tag. The samples are incubated for 5 minutes to one hour. Control samples are incubated in the absence of a test compound.

The buffer solution containing the test compounds is washed from the wells. Binding of a test compound to a PTPase-like protein polypeptide is detected by fluorescence measurements of the contents of the wells. A test compound that increases the fluorescence in a well by at least 15% relative to fluorescence of a well in which a test compound is not incubated is identified as a compound which binds to a PTPase-like protein polypeptide.

EXAMPLE 4

Identification of a test compound which decreases PTPase-like protein gene expression

A test compound is administered to a culture of human cells transfected with a PTPase-like protein expression construct and incubated at 37°C for 10 to 45 minutes. A culture of the same type of cells that have not been transfected is incubated for the same time without the test compound to provide a negative control.

RNA is isolated from the two cultures as described in Chirgwin *et al.*, *Biochem. 18*, 5294-99, 1979). Northern blots are prepared using 20 to 30 μg total RNA and hybridized with a ^{32}P-labeled PTPase-like specific-specific probe at 65°C in Express-hyb (CLONTECH). The probe comprises at least 11 contiguous nucleotides selected from the complement of SEQ ID NO: 1. A test compound that decreases the PTPase-like specific-specific signal relative to the signal obtained in the absence of the test compound is identified as an inhibitor of PTPase-like protein gene expression.

EXAMPLE 5

Identification of a test compound which decreases PTPase-like protein activity

A test compound is administered to a culture of human cells transfected with a PTPase-like protein expression construct and incubated at 37°C for 10 to 45 minutes. A culture of the same type of cells that have not been transfected is incubated for the same time without the test compound to provide a negative control. PTPase-like protein activity is measured using the method of Calvert-Evers & Hammond, *Cell Biol. Int. 24*, 559-68, 2000.

A test compound which decreases the PTPase-like protein activity of the PTPase-like protein relative to the PTPase-like protein activity in the absence of the test compound is identified as an inhibitor of PTPase-like protein activity.

EXAMPLE 6

Tissue-specific expression of PTPase-like protein

The qualitative expression pattern of PTPase-like protein in various tissues is determined by Reverse Transcription-Polymerase Chain Reaction (RT-PCR). To demonstrate that PTPase-like protein is involved in the disease process of obesity, expression is determined in the following tissues: subcutaneous adipose tissue, mesenteric adipose tissue, adrenal gland, bone marrow, brain (cerebellum, spinal cord, cerebral cortex, caudate, medulla, substantia nigra, and putamen), colon, fetal brain, heart, kidney, liver, lung, mammary gland, pancreas, placenta, prostate, salivary gland, skeletal muscle small intestine, spleen, stomach, testes, thymus, thyroid trachea, and uterus. Neuroblastoma cell lines SK-Nr-Be (2), Hr, Sk-N-As, HTB-10, IMR-32, SNSY-5Y, T3, SK-N-D2, D283, DAOY, CHP-2, U87MG, BE(2)C, T986, KANTS, MO59K, CHP234, C6 (rat), SK-N-F1, SK-PU-DW, PFSK-1, BE(2)M17, and MCIXC also are tested for PTPase-like protein expression. As a final step, the expression of PTPase-like protein in cells derived from normal individuals with the expression of cells derived from obese individuals is compared.

To demonstrate that PTPase-like protein is involved in the disease process of diabetes, the following whole body panel is screened to show predominant or relatively high expression: subcutaneous and mesenteric adipose tissue, adrenal gland, bone marrow, brain, colon, fetal brain, heart, hypothalamus, kidney, liver, lung, mammary gland, pancreas, placenta, prostate, salivary gland, skeletal muscle, small intestine, spleen, stomach, testis, thymus, thyroid, trachea, and uterus. Human islet cells and an islet cell library also are tested. As a final step, the expression of

PTPase-like protein in cells derived from normal individuals with the expression of cells derived from diabetic individuals is compared.

To demonstrate that PTPase-like protein is involved in CNS disorders, the following tissues are screened: fetal and adult brain, muscle, heart, lung, kidney, liver, thymus, testis, colon, placenta, trachea, pancreas, kidney, gastric mucosa, colon, liver, cerebellum, skin, cortex (Alzheimer's and normal), hypothalamus, cortex, amygdala, cerebellum, hippocampus, choroid, plexus, thalamus, and spinal cord.

To demonstrate that PTPase-like protein is involved in the disease process of COPD, the initial expression panel consists of RNA samples from respiratory tissues and inflammatory cells relevant to COPD: lung (adult and fetal), trachea, freshly isolated alveolar type II cells, cultured human bronchial epithelial cells, cultured small airway epithelial cells, cultured bronchial sooth muscle cells, cultured H441 cells (Clara-like), freshly isolated neutrophils and monocytes, and cultured monocytes (macrophage-like). Body map profiling also is carried out, using total RNA panels purchased from Clontech. The tissues are adrenal gland, bone marrow, brain, colon, heart, kidney, liver, lung, mammary gland, pancreas, prostate, salivary gland, skeletal muscle, small intestine, spleen, stomach, testis, thymus, trachea, thyroid, and uterus.

Quantitative expression profiling. Quantitative expression profiling is performed by the form of quantitative PCR analysis called "kinetic analysis" firstly described in Higuchi *et al., BioTechnology 10*, 413-17, 1992, and Higuchi *et al., BioTechnology 11*, 1026-30, 1993. The principle is that at any given cycle within the exponential phase of PCR, the amount of product is proportional to the initial number of template copies.

If the amplification is performed in the presence of an internally quenched fluorescent oligonucleotide (TaqMan probe) complementary to the target sequence, the probe is cleaved by the 5'-3' endonuclease activity of Taq DNA polymerase and a fluorescent dye released in the medium (Holland *et al., Proc. Natl. Acad. Sci. U.S.A.*

88, 7276-80, 1991). Because the fluorescence emission will increase in direct proportion to the amount of the specific amplified product, the exponential growth phase of PCR product can be detected and used to determine the initial template concentration (Heid *et al.*, *Genome Res. 6*, 986-94, 1996, and Gibson *et al.*, *Genome Res. 6*, 995-1001, 1996).

The amplification of an endogenous control can be performed to standardize the amount of sample RNA added to a reaction. In this kind of experiment, the control of choice is the 18S ribosomal RNA. Because reporter dyes with differing emission spectra are available, the target and the endogenous control can be independently quantified in the same tube if probes labeled with different dyes are used.

All "real time PCR" measurements of fluorescence are made in the ABI Prism 7700.

RNA extraction and cDNA preparation. Total RNA from the tissues listed above are used for expression quantification. RNAs labeled "from autopsy" were extracted from autoptic tissues with the TRIzol reagent (Life Technologies, MD) according to the manufacturer's protocol.

Fifty µg of each RNA were treated with DNase I for 1 hour at 37°C in the following reaction mix: 0.2 U/µl RNase-free DNase I (Roche Diagnostics, Germany); 0.4 U/µl RNase inhibitor (PE Applied Biosystems, CA); 10 mM Tris-HCl pH 7.9; 10mM $MgCl_2$; 50 mM NaCl; and 1 mM DTT.

After incubation, RNA is extracted once with 1 volume of phenol:chloroform:iso-amyl alcohol (24:24:1) and once with chloroform, and precipitated with 1/10 volume of 3 M NaAcetate, pH5.2, and 2 volumes of ethanol.

Fifty µg of each RNA from the autoptic tissues are DNase treated with the DNA-free kit purchased from Ambion (Ambion, TX). After resuspension and spectrophotometric quantification, each sample is reverse transcribed with the

TaqMan Reverse Transcription Reagents (PE Applied Biosystems, CA) according to the manufacturer's protocol. The final concentration of RNA in the reaction mix is 200ng/μL. Reverse transcription is carried out with 2.5μM of random hexamer primers.

TaqMan quantitative analysis. Specific primers and probe are designed according to the recommendations of PE Applied Biosystems; the probe can be labeled at the 5' end FAM (6-carboxy-fluorescein) and at the 3' end with TAMRA (6-carboxy-tetramethyl-rhodamine). Quantification experiments are performed on 10 ng of reverse transcribed RNA from each sample. Each determination is done in triplicate.

Total cDNA content is normalized with the simultaneous quantification (multiplex PCR) of the 18S ribosomal RNA using the Pre-Developed TaqMan Assay Reagents (PDAR) Control Kit (PE Applied Biosystems, CA).

The assay reaction mix is as follows: 1X final TaqMan Universal PCR Master Mix (from 2X stock) (PE Applied Biosystems, CA); 1X PDAR control – 18S RNA (from 20X stock); 300 nM forward primer; 900 nM reverse primer; 200 nM probe; 10 ng cDNA; and water to 25 μl.

Each of the following steps are carried out once: pre PCR, 2 minutes at 50°C, and 10 minutes at 95°C. The following steps are carried out 40 times: denaturation, 15 seconds at 95°C, annealing/extension, 1 minute at 60°C.

The experiment is performed on an ABI Prism 7700 Sequence Detector (PE Applied Biosystems, CA). At the end of the run, fluorescence data acquired during PCR are processed as described in the ABI Prism 7700 user's manual in order to achieve better background subtraction as well as signal linearity with the starting target quantity.

EXAMPLE 7

Diabetes: In vivo testing of compounds/target validation

1. Glucose Production:

Over-production of glucose by the liver, due to an enhanced rate of gluconeogenesis, is the major cause of fasting hyperglycemia in diabetes. Overnight fasted normal rats or mice have elevated rates of gluconeogenesis as do streptozotocin-induced diabetic rats or mice fed ad libitum. Rats are made diabetic with a single intravenous injection of 40 mg/kg of streptozotocin while C57BL/KsJ mice are given 40-60 mg/kg i.p. for 5 consecutive days. Blood glucose is measured from tail-tip blood and then compounds are administered via different routes (p.o., i.p., i.v., s.c.). Blood is collected at various times thereafter and glucose measured. Alternatively, compounds are administered for several days, then the animals are fasted overnight, blood is collected and plasma glucose measured. Compounds that inhibit glucose production will decrease plasma glucose levels compared to the vehicle-treated control group.

2. Insulin Sensitivity:

Both ob/ob and db/db mice as well as diabetic Zucker rats are hyperglycemic, hyperinsulinemic and insulin resistant. The animals are pre-bled, their glucose levels measured, and then they are grouped so that the mean glucose level is the same for each group. Compounds are administered daily either q.d. or b.i.d. by different routes (p.o., i.p., s.c.) for 7-28 days. Blood is collected at various times and plasma glucose and insulin levels determined. Compounds that improve insulin sensitivity in these models will decrease both plasma glucose and insulin levels when compared to the vehicle-treated control group.

3. Insulin Secretion:

Compounds that enhance insulin secretion from the pancreas will increase plasma insulin levels and improve the disappearance of plasma glucose following the administration of a glucose load. When measuring insulin levels, compounds are administered by different routes (p.o., i.p., s.c. or i.v.) to overnight fasted normal rats or mice. At the appropriate time an intravenous glucose load (0.4g/kg) is given, blood is collected one minute later. Plasma insulin levels are determined. Compounds that enhance insulin secretion will increase plasma insulin levels compared to animals given only glucose. When measuring glucose disappearance, animals are bled at the appropriate time after compound administration, then given either an oral or intraperitoneal glucose load (1g/kg), bled again after 15, 30, 60 and 90 minutes and plasma glucose levels determined. Compounds that increase insulin levels will decrease glucose levels and the area-under-the glucose curve when compared to the vehicle-treated group given only glucose.

EXAMPLE 8

In vivo testing of compounds/target validation

1. Pain:

Acute Pain

Acute pain is measured on a hot plate mainly in rats. Two variants of hot plate testing are used: In the classical variant animals are put on a hot surface (52 to 56 °C) and the latency time is measured until the animals show nocifensive behavior, such as stepping or foot licking. The other variant is an increasing temperature hot plate where the experimental animals are put on a surface of neutral temperature. Subsequently this surface is slowly but constantly heated until the animals begin to lick a hind paw. The temperature which is reached when hind paw licking begins is a measure for pain threshold.

Compounds are tested against a vehicle treated control group. Substance application is performed at different time points via different application routes (i.v., i.p., p.o., i.t., i.c.v., s.c., intradermal, transdermal) prior to pain testing.

Persistent Pain

Persistent pain is measured with the formalin or capsaicin test, mainly in rats. A solution of 1 to 5% formalin or 10 to 100 μg capsaicin is injected into one hind paw of the experimental animal. After formalin or capsaicin application the animals show nocifensive reactions like flinching, licking and biting of the affected paw. The number of nocifensive reactions within a time frame of up to 90 minutes is a measure for intensity of pain.

Compounds are tested against a vehicle treated control group. Substance application is performed at different time points via different application routes (i.v., i.p., p.o., i.t., i.c.v., s.c., intradermal, transdermal) prior to formalin or capsaicin administration.

Neuropathic Pain

Neuropathic pain is induced by different variants of unilateral sciatic nerve injury mainly in rats. The operation is performed under anesthesia. The first variant of sciatic nerve injury is produced by placing loosely constrictive ligatures around the common sciatic nerve. The second variant is the tight ligation of about the half of the diameter of the common sciatic nerve. In the next variant, a group of models is used in which tight ligations or transections are made of either the L5 and L6 spinal nerves, or the L% spinal nerve only. The fourth variant involves an axotomy of two of the three terminal branches of the sciatic nerve (tibial and common peroneal nerves) leaving the remaining sural nerve intact whereas the last variant comprises the axotomy of only the tibial branch leaving the sural and common nerves uninjured. Control animals are treated with a sham operation.

Postoperatively, the nerve injured animals develop a chronic mechanical allodynia,

cold allodynioa, as well as a thermal hyperalgesia. Mechanical allodynia is measured by means of a pressure transducer (electronic von Frey Anesthesiometer, IITC Inc.-Life Science Instruments, Woodland Hills, SA, USA; Electronic von Frey System, Somedic Sales AB, Hörby, Sweden). Thermal hyperalgesia is measured by means of a radiant heat source (Plantar Test, Ugo Basile, Comerio, Italy), or by means of a cold plate of 5 to 10 °C where the nocifensive reactions of the affected hind paw are counted as a measure of pain intensity. A further test for cold induced pain is the counting of nocifensive reactions, or duration of nocifensive responses after plantar administration of acetone to the affected hind limb. Chronic pain in general is assessed by registering the circadanian rhythms in activity (Surjo and Arndt, Universität zu Köln, Cologne, Germany), and by scoring differences in gait (foot print patterns; FOOTPRINTS program, Klapdor et al., 1997. A low cost method to analyze footprint patterns. J. Neurosci. Methods 75, 49-54).

Compounds are tested against sham operated and vehicle treated control groups. Substance application is performed at different time points via different application routes (i.v., i.p., p.o., i.t., i.c.v., s.c., intradermal, transdermal) prior to pain testing.

Inflammatory Pain

Inflammatory pain is induced mainly in rats by injection of 0.75 mg carrageenan or complete Freund's adjuvant into one hind paw. The animals develop an edema with mechanical allodynia as well as thermal hyperalgesia. Mechanical allodynia is measured by means of a pressure transducer (electronic von Frey Anesthesiometer, IITC Inc.-Life Science Instruments, Woodland Hills, SA, USA). Thermal hyperalgesia is measured by means of a radiant heat source (Plantar Test, Ugo Basile, Comerio, Italy, Paw thermal stimulator, G. Ozaki, University of California, USA). For edema measurement two methods are being used. In the first method, the animals are sacrificed and the affected hindpaws sectioned and weighed. The second method comprises differences in paw volume by measuring water displacement in a plethysmometer (Ugo Basile, Comerio, Italy).

Compounds are tested against uninflamed as well as vehicle treated control groups. Substance application is performed at different time points via different application routes (i.v., i.p., p.o., i.t., i.c.v., s.c., intradermal, transdermal) prior to pain testing.

Diabetic Neuropathic Pain

Rats treated with a single intraperitoneal injection of 50 to 80 mg/kg streptozotocin develop a profound hyperglycemia and mechanical allodynia within 1 to 3 weeks. Mechanical allodynia is measured by means of a pressure transducer (electronic von Frey Anesthesiometer, IITC Inc.-Life Science Instruments, Woodland Hills, SA, USA).

Compounds are tested against diabetic and non-diabetic vehicle treated control groups. Substance application is performed at different time points via different application routes (i.v., i.p., p.o., i.t., i.c.v., s.c., intradermal, transdermal) prior to pain testing.

2. Parkinson's disease

6-Hydroxydopamine (6-OH-DA) Lesion

Degeneration of the dopaminergic nigrostriatal and striatopallidal pathways is the central pathological event in Parkinson's disease. This disorder has been mimicked experimentally in rats using single/sequential unilateral stereotaxic injections of 6-OH-DA into the medium forebrain bundle (MFB).

Male Wistar rats (Harlan Winkelmann, Germany), weighing 200±250 g at the beginning of the experiment, are used. The rats are maintained in a temperature- and humidity-controlled environment under a 12 h light/dark cycle with free access to food and water when not in experimental sessions. The following in vivo protocols are approved by the governmental authorities. All efforts are made to minimize animal suffering, to reduce the number of animals used, and to utilize alternatives to in vivo techniques.

Animals are administered pargyline on the day of surgery (Sigma, St. Louis, MO, USA; 50 mg/kg i.p.) in order to inhibit metabolism of 6-OHDA by monoamine oxidase and desmethylimipramine HCl (Sigma; 25 mg/kg i.p.) in order to prevent uptake of 6-OHDA by noradrenergic terminals. Thirty minutes later the rats are anesthetized with sodium pentobarbital (50 mg/kg) and placed in a stereotaxic frame. In order to lesion the DA nigrostriatal pathway 4 μl of 0.01% ascorbic acid-saline containing 8 μg of 6-OHDA HBr (Sigma) are injected into the left medial fore-brain bundle at a rate of 1 μl/min (2.4 mm anterior, 1.49 mm lateral, -2.7 mm ventral to Bregma and the skull surface). The needle is left in place an additional 5 min to allow diffusion to occur.

Stepping Test

Forelimb akinesia is assessed three weeks following lesion placement using a modified stepping test protocol. In brief, the animals are held by the experimenter with one hand fixing the hindlimbs and slightly raising the hind part above the surface. One paw is touching the table, and is then moved slowly sideways (5 s for 1 m), first in the forehand and then in the backhand direction. The number of adjusting steps is counted for both paws in the backhand and forehand direction of movement. The sequence of testing is right paw forehand and backhand adjusting stepping, followed by left paw forehand and backhand directions. The test is repeated three times on three consecutive days, after an initial training period of three days prior to the first testing. Forehand adjusted stepping reveals no consistent differences between lesioned and healthy control animals. Analysis is therefore restricted to backhand adjusted stepping.

Balance Test

Balance adjustments following postural challenge are also measured during the stepping test sessions. The rats are held in the same position as described in the stepping test and, instead of being moved sideways, tilted by the experimenter towards the side of the paw touching the table. This maneuver results in loss of

balance and the ability of the rats to regain balance by forelimb movements is scored on a scale ranging from 0 to 3. Score 0 is given for a normal forelimb placement. When the forelimb movement is delayed but recovery of postural balance detected, score 1 is given. Score 2 represents a clear, yet insufficient, forelimb reaction, as evidenced by muscle contraction, but lack of success in recovering balance, and score 3 is given for no reaction of movement. The test is repeated three times a day on each side for three consecutive days after an initial training period of three days prior to the first testing.

Staircase Test (Paw Reaching)

A modified version of the staircase test is used for evaluation of paw reaching behavior three weeks following primary and secondary lesion placement. Plexiglass test boxes with a central platform and a removable staircase on each side are used. The apparatus is designed such that only the paw on the same side at each staircase can be used, thus providing a measure of independent forelimb use. For each test the animals are left in the test boxes for 15 min. The double staircase is filled with 7 x 3 chow pellets (Precision food pellets, formula: P, purified rodent diet, size 45 mg; Sandown Scientific) on each side. After each test the number of pellets eaten (successfully retrieved pellets) and the number of pellets taken (touched but dropped) for each paw and the success rate (pellets eaten/pellets taken) are counted separately. After three days of food deprivation (12 g per animal per day) the animals are tested for 11 days. Full analysis is conducted only for the last five days.

MPTP treatment

The neurotoxin 1-methyl-4-phenyl-1,2,3,6-tetrahydro-pyridine (MPTP) causes degeneration of mesencephalic dopaminergic (DAergic) neurons in rodents, non-human primates, and humans and, in so doing, reproduces many of the symptoms of Parkinson's disease. MPTP leads to a marked decrease in the levels of dopamine and its metabolites, and in the number of dopaminergic terminals in the striatum as well as severe loss of the tyrosine hydroxylase (TH)-immunoreactive cell bodies in the substantia nigra, pars compacta.

In order to obtain severe and long-lasting lesions, and to reduce mortality, animals receive single injections of MPTP, and are then tested for severity of lesion 7–10 days later. Successive MPTP injections are administered on days 1, 2 and 3. Animals receive application of 4 mg/kg MPTP hydrochloride (Sigma) in saline once daily. All injections are intraperitoneal (i.p.) and the MPTP stock solution is frozen between injections. Animals are decapitated on day 11.

Immunohistology

At the completion of behavioral experiments, all animals are anaesthetized with 3 ml thiopental (1 g/40 ml i.p., Tyrol Pharma). The mice are perfused transcardially with 0.01 M PBS (pH 7.4) for 2 min, followed by 4% paraformaldehyde (Merck) in PBS for 15 min. The brains are removed and placed in 4% paraformaldehyde for 24 h at 4 °C. For dehydration they are then transferred to a 20% sucrose (Merck) solution in 0.1 M PBS at 4 °C until they sink. The brains are frozen in methylbutan at -20 °C for 2 min and stored at -70 °C. Using a sledge microtome (mod. 3800-Frigocut, Leica), 25 μm sections are taken from the genu of the corpus callosum (AP 1.7 mm) to the hippocampus (AP 21.8 mm) and from AP 24.16 to AP 26.72. Forty-six sections are cut and stored in assorters in 0.25 M Tris buffer (pH 7.4) for immunohistochemistry.

A series of sections is processed for free-floating tyrosine hydroxylase (TH) immunohistochemistry. Following three rinses in 0.1 M PBS, endogenous peroxidase activity is quenched for 10 min in 0.3% H_2O_2 ±PBS. After rinsing in PBS, sections are preincubated in 10% normal bovine serum (Sigma) for 5 min as blocking agent and transferred to either primary anti-rat TH rabbit antiserum (dilution 1:2000).

Following overnight incubation at room temperature, sections for TH immunoreactivity are rinsed in PBS (2 x10 min) and incubated in biotinylated anti-rabbit immunoglobulin G raised in goat (dilution 1:200) (Vector) for 90 min, rinsed repeatedly and transferred to Vectastain ABC (Vector) solution for 1 h. 3,.3' -Diaminobenzidine tetrahydrochloride (DAB; Sigma) in 0.1 M PBS, supplemented

with 0.005% H_2O_2 , serves as chromogen in the subsequent visualization reaction. Sections are mounted on to gelatin-coated slides, left to dry overnight, counter-stained with hematoxylin dehydrated in ascending alcohol concentrations and cleared in butylacetate. Coverslips are mounted on entellan.

Rotarod Test

We use a modification of the procedure described by Rozas and Labandeira-Garcia (1997), with a CR-1 Rotamex system (Columbus Instruments, Columbus, OH) comprising an IBM-compatible personal computer, a CIO-24 data acquisition card, a control unit, and a four-lane rotarod unit. The rotarod unit consists of a rotating spindle (diameter 7.3 cm) and individual compartments for each mouse. The system software allows preprogramming of session protocols with varying rotational speeds (0–80 rpm). Infrared beams are used to detect when a mouse has fallen onto the base grid beneath the rotarod. The system logs the fall as the end of the experiment for that mouse, and the total time on the rotarod, as well as the time of the fall and all the set-up parameters, are recorded. The system also allows a weak current to be passed through the base grid, to aid training.

3. Dementia

The object recognition task

The object recognition task has been designed to assess the effects of experimental manipulations on the cognitive performance of rodents. A rat is placed in an open field, in which two identical objects are present. The rats inspects both objects during the first trial of the object recognition task. In a second trial, after a retention interval of for example 24 hours, one of the two objects used in the first trial, the 'familiar' object, and a novel object are placed in the open field. The inspection time at each of the objects is registered. The basic measures in the OR task is the time spent by a rat exploring the two object the second trial. Good retention is reflected by higher exploration times towards the novel than the 'familiar' object.

Administration of the putative cognition enhancer prior to the first trial predominantly allows assessment of the effects on acquisition, and eventually on consolidation processes. Administration of the testing compound after the first trial allows to assess the effects on consolidation processes, whereas administration before the second trial allows to measure effects on retrieval processes.

The passive avoidance task

The passive avoidance task assesses memory performance in rats and mice. The inhibitory avoidance apparatus consists of a two-compartment box with a light compartment and a dark compartment. The two compartments are separated by a guillotine door that can be operated by the experimenter. A threshold of 2 cm separates the two compartments when the guillotine door is raised. When the door is open, the illumination in the dark compartment is about 2 lux. The light intensity is about 500 lux at the center of the floor of the light compartment.

Two habituation sessions, one shock session, and a retention session are given, separated by inter-session intervals of 24 hours. In the habituation sessions and the retention session the rat is allowed to explore the apparatus for 300 sec. The rat is placed in the light compartment, facing the wall opposite to the guillotine door. After an accommodation period of 15 sec. the guillotine door is opened so that all parts of the apparatus can be visited freely. Rats normally avoid brightly lit areas and will enter the dark compartment within a few seconds.

In the shock session the guillotine door between the compartments is lowered as soon as the rat has entered the dark compartment with its four paws, and a scrambled 1 mA footshock is administered for 2 sec. The rat is removed from the apparatus and put back into its home cage. The procedure during the retention session is identical to that of the habituation sessions.

The step-through latency, that is the first latency of entering the dark compartment (in sec.) during the retention session is an index of the memory performance of the

animal; the longer the latency to enter the dark compartment, the better the retention is. A testing compound in given half an hour before the shock session, together with 1 $mg*kg^{-1}$ scopolamine. Scopolamine impairs the memory performance during the retention session 24 hours later. If the test compound increases the enter latency compared with the scopolamine-treated controls, is likely to possess cognition enhancing potential.

The Morris water escape task

The Morris water escape task measures spatial orientation learning in rodents. It is a test system that has extensively been used to investigate the effects of putative therapeutic on the cognitive functions of rats and mice. The performance of an animal is assessed in a circular water tank with an escape platform that is submerged about 1 cm below the surface of the water. The escape platform is not visible for an animal swimming in the water tank. Abundant extra-maze cues are provided by the furniture in the room, including desks, computer equipment, a second water tank, the presence of the experimenter, and by a radio on a shelf that is playing softly.

The animals receive four trials during five daily acquisition sessions. A trial is started by placing an animal into the pool, facing the wall of the tank. Each of four starting positions in the quadrants north, east, south, and west is used once in a series of four trials; their order is randomized. The escape platform is always in the same position. A trial is terminated as soon as the animal had climbs onto the escape platform or when 90 seconds have elapsed, whichever event occurs first. The animal is allowed to stay on the platform for 30 seconds. Then it is taken from the platform and the next trial is started. If an animal did not find the platform within 90 seconds it is put on the platform by the experimenter and is allowed to stay there for 30 seconds. After the fourth trial of the fifth daily session, an additional trial is given as a probe trial: the platform is removed, and the time the animal spends in the four quadrants is measured for 30 or 60 seconds. In the probe trial, all animals start from the same start position, opposite to the quadrant where the escape platform had been positioned during acquisition.

Four different measures are taken to evaluate the performance of an animal during acquisition training: escape latency, traveled distance, distance to platform, and swimming speed. The following measures are evaluated for the probe trial: time (s) in quadrants and traveled distance (cm) in the four quadrants. The probe trial provides additional information about how well an animal learned the position of the escape platform. If an animal spends more time and swims a longer distance in the quadrant where the platform had been positioned during the acquisition sessions than in any other quadrant, one concludes that the platform position has been learned well.

In order to assess the effects of putative cognition enhancing compounds, rats or mice with specific brain lesions which impair cognitive functions, or animals treated with compounds such as scopolamine or MK-801, which interfere with normal learning, or aged animals which suffer from cognitive deficits, are used.

The T-maze spontaneous alternation task

The T-maze spontaneous alternation task (TeMCAT) assesses the spatial memory performance in mice. The start arm and the two goal arms of the T-maze are provided with guillotine doors which can be operated manually by the experimenter. A mouse is put into the start arm at the beginning of training. The guillotine door is closed. In the first trial, the 'forced trial', either the left or right goal arm is blocked by lowering the guillotine door. After the mouse has been released from the start arm, it will negotiate the maze, eventually enter the open goal arm, and return to the start position, where it will be confined for 5 seconds, by lowering the guillotine door. Then, the animal can choose freely between the left and right goal arm (all guillotine-doors opened) during 14 'free choice' trials. As soon a the mouse has entered one goal arm, the other one is closed. The mouse eventually returns to the start arm and is free to visit whichever go alarm it wants after having been confined to the start arm for 5 seconds. After completion of 14 free choice trials in one session, the animal is removed from the maze. During training, the animal is never handled.

The percent alternations out of 14 trials is calculated. This percentage and the total time needed to complete the first forced trial and the subsequent 14 free choice trials (in s) is analyzed. Cognitive deficits are usually induced by an injection of scopolamine, 30 min before the start of the training session. Scopolamine reduced the per-cent alternations to chance level, or below. A cognition enhancer, which is always administered before the training session, will at least partially, antagonize the scopolamine-induced reduction in the spontaneous alternation rate.

EXAMPLE 9

Treatment of COPD in an animal model

Guinea pigs are exposed on a single occasion to tobacco smoke for 50 minutes. Animals are sacrificed between 10 minutes and 24 hour following the end of the exposure and their lungs placed in RNAlater™. The lung tissue is homogenized and total RNA is extracted using a Qiagens RNeasy™ Maxi kit. Molecular Probes RiboGreen™ RNA quantitation method is used to quantify the amount of RNA in each sample. Total RNA is reverse transcribed and the resultant cDNA was used in a real-time polymerase chain reaction (PCR). The cDNA is added to a solution containing the sense and anti-sense primers and the 6-carboxy-tetramethyl-rhodamine labeled probe of the "COPD GENE". Cyclophilin is used as the housekeeping gene. The expression of the "COPD GENE" is measured using the TaqMan real-time PCR system that generates an amplification curve for each sample. From this curve a threshold cycle value is calculated: the fractional cycle number at which the amount of amplified target reaches a fixed threshold. A sample containing many copies of the "COPD GENE" will reach this threshold earlier than a sample containing fewer copies. The threshold is set at 0.2 and the threshold cycle C_T is calculated from the amplification curve. The C_T value for the "COPD GENE" is normalized using the C_T value for the housekeeping gene.

Expression of the "COPD GENE" is increased by at least 1,5-fold between 10 minutes and 3 hours post tobacco smoke exposure compared to air exposed control animals.

Test compounds are evaluated as follows. Animals are pre-treated with a test compound between 5 minutes and 1 hour prior to the tobacco smoke exposure and they are then sacrificed up to 3 hours after the tobacco smoke exposure has been completed. Control animals are pre-treated with the vehicle of the test compound via the route of administration chosen for the test compound. A test compound that reduces the tobacco smoke induced upregulation of the "COPD GENE" relative to the expression seen in vehicle treated tobacco smoke exposed animals is identified as an inhibitor of "COPD GENE" expression.

REFERENCES

1. Structure and function of the protein tyrosine phosphatases. Trends Biochem Sci 1996 Nov;21(11):413-7.

2. Crystal structure of the MAPK phosphatase Pyst1 catalytic domain and implications for regulated activation. Nat Struct Biol 1999 Feb;6(2):174-81.

3. Protein tyrosine phosphatases in signal transduction. Curr Opin Cell Biol 1997 Apr;9(2):193-204.

CLAIMS

1. An isolated polynucleotide being selected from the group consisting of:
   a. a polynucleotide encoding a PTPase-like protein polypeptide comprising an amino acid sequence selected form the group consisting of:
      i. amino acid sequences which are at least about 95% identical to the amino acid sequence shown in SEQ ID NO: 2; and
      ii. the amino acid sequence shown in SEQ ID NO: 2.
   b. a polynucleotide comprising the sequence of SEQ ID NO: 1;
   c. a polynucleotide which hybridizes under stringent conditions to a polynucleotide specified in (a) and (b) and encodes a PTPase-like protein polypeptide;
   d. a polynucleotide the sequence of which deviates from the polynucleotide sequences specified in (a) to (c) due to the degeneration of the genetic code and encodes a PTPase-like protein polypeptide; and
   e. a polynucleotide which represents a fragment, derivative or allelic variation of a polynucleotide sequence specified in (a) to (d) and encodes a PTPase-like protein polypeptide.

2. An expression vector containing any polynucleotide of claim 1.

3. A host cell containing the expression vector of claim 2.

4. A substantially purified PTPase-like protein polypeptide encoded by a polynucleotide of claim 1.

5. A method for producing a PTPase-like protein polypeptide, wherein the method comprises the following steps:

a. culturing the host cell of claim 3 under conditions suitable for the expression of the PTPase-like protein polypeptide; and
b. recovering the PTPase-like protein polypeptide from the host cell culture.

6. A method for detection of a polynucleotide encoding a PTPase-like protein polypeptide in a biological sample comprising the following steps:
 a. hybridizing any polynucleotide of claim 1 to a nucleic acid material of a biological sample, thereby forming a hybridization complex; and
 b. detecting said hybridization complex.

7. The method of claim 6, wherein before hybridization, the nucleic acid material of the biological sample is amplified.

8. A method for the detection of a polynucleotide of claim 1 or a PTPase-like protein polypeptide of claim 4 comprising the steps of:
 a. contacting a biological sample with a reagent which specifically interacts with the polynucleotide or the PTPase-like protein polypeptide; and
 b. detecting the interaction.

9. A diagnostic kit for conducting the method of any one of claims 6 to 8.

10. A method of screening for agents which decrease the activity of a PTPase-like protein, comprising the steps of:
 a. contacting a test compound with any PTPase-like protein polypeptide encoded by any polynucleotide of claim1;
 b. detecting binding of the test compound to the PTPase-like protein polypeptide, wherein a test compound which binds to the polypeptide is identified as a potential therapeutic agent for decreasing the activity of a PTPase-like protein.

11. A method of screening for agents which regulate the activity of a PTPase-like protein, comprising the steps of:
    a. contacting a test compound with a PTPase-like protein polypeptide encoded by any polynucleotide of claim 1; and
    b. detecting a PTPase-like protein activity of the polypeptide, wherein a test compound which increases the PTPase-like protein activity is identified as a potential therapeutic agent for increasing the activity of the PTPase-like protein, and wherein a test compound which decreases the PTPase-like protein activity of the polypeptide is identified as a potential therapeutic agent for decreasing the activity of the PTPase-like protein.

12. A method of screening for agents which decrease the activity of a PTPase-like protein, comprising the steps of:
    a. contacting a test compound with any polynucleotide of claim 1;
    b. and detecting binding of the test compound to the polynucleotide, wherein a test compound which binds to the polynucleotide is identified as a potential therapeutic agent for decreasing the activity of PTPase-like protein.

13. A method of reducing the activity of PTPase-like protein, comprising the step of contacting a cell with a reagent which specifically binds to any polynucleotide of claim 1 or any PTPase-like protein polypeptide of claim 4, whereby the activity of PTPase-like protein is reduced.

14. A reagent that modulates the activity of a PTPase-like protein polypeptide or a polynucleotide wherein said reagent is identified by the method of any of the claim 10 to 12.

15. A pharmaceutical composition, comprising the expression vector of claim 2 or the reagent of claim 14 and a pharmaceutically acceptable carrier.

16. Use of the expression vector of claim 2 or the reagent of claim 14 in the preparation of a medicament for modulating the activity of a PTPase-like protein in a disease.

17. Use of claim 16 wherein the disease is obesity, diabetes, a CNS disorder, a cardiovascular disorder or COPD.

Fig. 1

```
atggcatgttccatggcgtgtggcggtagagcttgcaagtatgagaac
ccagcccgctggagtgagcaggagcaagccattaaggggtttactca
tcctgggtcactgataatatactggccatggcccgcccatcctctgag
ctcctggagaagtaccacatcattgatcagttcctcagccatggcata
aaaacaataatcaacctccagcgccctggtgagcatgctagctgtggg
aaccctctggaacaagaaagtggcttcacataccttcctgaggctttc
atggaggctggcatttacttctacaatttcggatggaaggattatggt
gtagcgtctcttactactatcctagatatggtgaaggtgatgacattt
gccttacaggaaggaaaagtagctatccattgtcatgcagggcttggt
cgaacaggtgttttaatagcctgttacttagtttttgcaacgagaatg
actgctgaccaagcaattatatttgtgcgggcaaagcgacccaattcc
atacaaaccagaggacagctcctctgtgtaagggaatttactcagttt
ctaactcctctccgcaatatattctcttgctgtgatcccaaagcacat
gctgtcaccttacctcaatatctaattcgccagcgtcatctgcttcat
ggttatgaggcacgacttctgaaacacgtgccaaaaattatccaccta
gtttgcaaattgctgctggacttagcggagaacaggccagtgatgatg
aaggatgtgtccgaaggacctggtctctctgctgaaatagaaaagaca
atgtctgagatggtcaccatgcagctggataaagagttactgaggcat
gacagtgatgtgtccaacccgcctaaccccactgcagtggcagcagat
tttgacaatcgaggcatgatttctccaatgagcaacagtttgaccct
ctttggaaaaggcggaatgttgagtgccttcaaccccctgactcatctg
aaaaggcggctcagctacagtgactcagatttaaagagggccgagaac
ctcctggagcaaggggagactccacagacagtgcctgcccagatcttg
gttggccacaagcccaggcagcagaagctcataagccattgttacatc
ccacagtctccagaaccagacttacacaaggaagccttggttcgcagc
acactttctttctggagtcagtcaaagtttggaggcctggaaggactc
aaagataatgggtcaccaattttccatggaaggatcattccaaaggaa
gcacagcagagtggagctttctctgcagatgtttcaggctcacacagc
cctggggagccagtttcacccagctttgcaaatgtccataaggatcca
aaccctgctcaccagcaagtgtctcactgtcagtgtaaaactcatggt
gttgggagccctggctctgtcaggcagaacagcaggacaccccgaagc
cctctggactgtggctccagtcccaaagcacagttcttggttgaacat
gaaacccaggacagtaaagatctgtctgaagcagcttcacactctgca
ttacagtctgaattgagtgctgaggcaagaagaatactggcggccaaa
gccctagcaaatttaaatgaatctgtagaaaaggaggaactaaaaagg
aagaaagagcttaattcccgagatggagcttgggaaagaatatgtggc
gagagggaccctttcatcctatgcagcttgatgtggtcttgggtggag
caactgaaggagcctgtaatcaccaaagaggatgtggacatgttggtt
gacaggcgagcagatgccgcagaagcactttttttattagagaaggga
cagcaccagactattctctgcgtgttgcactgcatagtgaacctgcag
acaattcccgtggatgtggaggaagctttccttgcccatgccattaag
gcattcactaaggttaattttgattctgaaaatggaccaacagtttac
aacaccctgaagaaaatatttaagcacacg
```

Fig. 2

MACSMACGGRACKYENPARWSEQEQAIKGVYSSWVTDNILAMARPSSE
LLEKYHIIDQFLSHGIKTIINLQRPGEHASCGNPLEQESGFTYLPEAF
MEAGIYFYNFGWKDYGVASLTTILDMVKVMTFALQEGKVAIHCHAGLG
RTGVLIACYLVFATRMTADQAIIFVRAKRPNSIQTRGQLLCVREFTQF
LTPLRNIFSCCDPKAHAVTLPQYLIRQRHLLHGYEARLLKHVPKIIHL
VCKLLLDLAENRPVMMKDVSEGPGLSAEIEKTMSEMVTMQLDKELLRH
DSDVSNPPNPTAVAADFDNRGMIFSNEQQFDPLWKRRNVECLQPLTHL
KRRLSYSDSDLKRAENLLEQGETPQTVPAQILVGHKPRQQKLISHCYI
PQSPEPDLHKEALVRSTLSFWSQSKFGGLEGLKDNGSPIFHGRIIPKE
AQQSGAFSADVSGSHSPGEPVSPSFANVHKDPNPAHQQVSHCQCKTHG
VGSPGSVRQNSRTPRSPLDCGSSPKAQFLVEHETQDSKDLSEAASHSA
LQSELSAEARRILAAKALANLNESVEKEELKRKKELNSRDGAWERICG
ERDPFILCSLMWSWVEQLKEPVITKEDVDMLVDRRADAAEALFLLEKG
QHQTILCVLHCIVNLQTIPVDVEEAFLAHAIKAFTKVNFDSENGPTVY
NTLKKIFKHT

Fig. 3

```
   1 aggagagcgt ttacggcatg tcattcctgg acacatggcg tgttccatgg cgtgtggtgg
  61 tagagcttgc aagtatgaga atccagcccg ctggagtgag caggaacaag ccattaaagg
 121 ggtttactca tcctgggtca ctgataatat actggccatg gcccgcccat cctctgagct
 181 cctggagaag taccacatca ttgatcagtt cctcagccat ggcataaaaa caataatcaa
 241 cctccagcgc cctggtgagc atgctagctg tgggaaccct ctggaacaag aaagtggctt
 301 cacatacctt cctgaggctt tcatggaggc tggcatttat ttctacaatt tcgggtggaa
 361 ggattatggt gtagcatctc ttactactat cctagatatg gtgaaggtga tgacatttgc
 421 cttacaggaa ggaaaagtag ctatccattg tcatgcaggg cttggtcgaa caggtgtctt
 481 aatagcatgt tacttagttt ttgcaacaag aatgactgct gaccaagcaa ttatatttgt
 541 gcgggcaaag cgacccaatt ccatacaaac cagagggcag ctcctctgtg taagggaatt
 601 tactcagttt ctaactcctc tccgcaatat attctcttgc tgtgatccca aagcacatgc
 661 tgtcacctta gctcaatatc taattcgcca gcgtcatctg cttcatggtt atgaggcacg
 721 acttctgaaa catgtgccaa aaattatcca cctagtttgc aaattgcttc tggacttagc
 781 ggaaaacagg ccagtgatga tgaaggacga gtccgaagga cctggtctct ctgctgaaat
 841 agaaaagacc atgtctgaga tggtcaccat gcagctggat aaagaattac tgaggaatga
 901 cagtgatgtg tccaacccgc ctaaccccac.tgcagtggca gcagattttg acaatcgagg
 961 catgatttc tccaatgagc aagagtttga ccctctttgg aaaaggcgga atgttgagtg
1021 ccttcaaccc ctttctcatc tgaaaaggcg gctcagctac agtgactcag atttaaagag
1081 ggccgagaac ctcctggagg aaggggagcc tccacagaca gtgcctgtcc aggtcttggt
1141 tggccacaac cccaggcagc agaagcgcat aagccattgt tacatcccac agtcttcaga
1201 accagactta cataaagaag ccttggttcg cagcacactt tctttctgga gtcagtctaa
1261 gtttggaagg ctggaagggc tcaaagataa tggctcacca attttccatg gacggatcat
1321 tccaaaggaa gcacagcaga gtagagcttt ctctgcagat gtttcaggcc cacacagccc
1381 tggggaacca gtttcaccca actttgcaaa tgtccgtaag gatccaaacc ctactcacca
1441 gcaagtgtct cactgtcagt gtaaaactca tggtgttggg agccctgggt ctgtcaggca
1501 gaacagcagg acaccccaaa gacctctgga ctgtggctgc agtctgaaag cacagttctt
1561 ggttgaacat gaaacccagg acagtaaaga tctggctgaa gcagcttcac actctacatt
1621 acagtctgaa ttgagtgctg aggcaagaag aatactggcg gccaaagctc tagcaaattt
1681 aaatgaatnt gtagaaaagg aggaactaaa aaggaagaaa gagcttaatt cccgagatgg
1741 agcttgggaa agaatatgtg gcgagaggga ccctttcatc ctatgcagct tgatgtggtc
1801 ttgggtggag caactgaagg agcctgtaat caccaaagag gatgtggaca tgttggttga
1861 caggcgagca gacgccgcag aagcactgtt tttattagag aagggacagc accagactat
1981 tttccttgcc catgccatta aggcgttcac taaggttaat tttgattctg aaaatggacc
2041 aatagtttac aacaccctga agaaaatatt taagcacacg cggtga
```

Fig. 4

BLASTP - alignment of 406 against trembl|AB046040|AB046040_1product:
"**unnamed** protein product"; Macaca fascicularis brain cDNA,clone:QccE-10114.
//:gp|AB046040|9280136 product: "unnamed protein product"; Macaca fascicularis brain cDNA, clone:QccE-10114.
This hit is scoring at : 0.0 (expectation value)
Alignment length (overlap) : 682
Identities : 94 %
Scoring matrix : BLOSUM62 (used to infer consensus pattern)
Database searched : nrdb_1_;

```
Q:      1 MACSMACGGRACKYENPARWSEQEQAIKGVYSSWVTDNILAMARPSSELLEKYHIIDQFL
          MACSMACGGRACKYENPARWSEQEQAIKGVYSSWVTDNILAMARPSSELLEKYHIIDQFL
H:      1 MACSMACGGRACKYENPARWSEQEQAIKGVYSSWVTDNILAMARPSSELLEKYHIIDQFL

          SHGIKTIINLQRPGEHASCGNPLEQESGFTYLPEAFMEAGIYFYNFGWKDYGVASLTTIL
          SHGIKTIINLQRPGEHASCGNPLEQESGFTYLPEAFMEAGIYFYNFGWKDYGVASLTTIL
          SHGIKTIINLQRPGEHASCGNPLEQESGFTYLPEAFMEAGIYFYNFGWKDYGVASLTTIL
                           active site
          DMVKVMTFALQEGKVAIHCHAGLGRTGVLIACYLVFATRMTADQAIIFVRAKRPNSIQTR
          DMVKVMTFALQEGKVAIHCHAGLGRTGVLIACYLVFATRMTADQAIIFVRAKRPNSIQTR
          DMVKVMTFALQEGKVAIHCHAGLGRTGVLIACYLVFATRMTADQAIIFVRAKRPNSIQTR

          GQLLCVREFTQFLTPLRNIFSCCDPKAHAVTLPQYLIRQRHLLHGYEARLLKHVPKIIHL
          GQLLCVREFTQFLTPLRNIFSCCDPKAHAVTL.QYLIRQRHLLHGYEARLLKHVPKIIHL
          GQLLCVREFTQFLTPLRNIFSCCDPKAHAVTLAQYLIRQRHLLHGYEARLLKHVPKIIHL

          VCKLLLDLAENRPVMMKDVSEGPGLSAEIEKTMSEMVTMQLDKELLRHDSDVSNPPNPTA
          VCKLLLDLAENRPVMMKD SEGPGLSAEIEKTMSEMVTMQLDKELLR:DSDVSNPPNPTA
          VCKLLLDLAENRPVMMKDESEGPGLSAEIEKTMSEMVTMQLDKELLRNDSDVSNPPNPTA

          VAADFDNRGMIFSNEQQFDPLWKRRNVECLQPLTHLKRRLSYSDSDLKRAENLLEQGETP
          VAADFDNRGMIFSNEQ:FDPLWKRRNVECLQPL:HLKRRLSYSDSDLKRAENLLE:GE.P
          VAADFDNRGMIFSNEQEFDPLWKRRNVECLQPLSHLKRRLSYSDSDLKRAENLLEEGEPP
```

Fig. 4 (cont'd)

```
QTVPAQILVGHKPRQQKLISHCYIPQSPEPDLHKEALVRSTLSFWSQSKFGGLEGLKDNG
QTVP.Q:LVGH.PRQQK ISHCYIPQS.EPDLHKEALVRSTLSFWSQSKFG LEGLKDNG
QTVPVQVLVGHNPRQQKRISHCYIPQSSEPDLHKEALVRSTLSFWSQSKFGRLEGLKDNG

SPIFHGRIIPKEAQQXGAFXADVXGXHXPGEPVXPXFANVHKDPNPAHQQVXHCQCKTHG
SPIFHGRIIPKEAQQ. AF.ADV.G H.PGEPV.P.FANV.KDPNP.HQQV.HCQCKTHG
SPIFHGRIIPKEAQQSRAFSADVSGPHSPGEPVSPNFANVRKDPNPTHQQVSHCQCKTHG

VGXPGXVRQNXRTPRXPLDCGXXPKAQFLVEHETQDXKDLXEAAXHXALQXELXAEARRI
VG.PG.VRQN.RTP:.PLDCG . KAQFLVEHETQD.KDL.EAA.H..LQ.EL.AEARRI
VGSPGSVRQNSRTPQRPLDCGCSLKAQFLVEHETQDSKDLAEAASHSTLQSELSAEARRI

LAAKALANLNEXVEKEELKRKKELNSRDGAWERICGERDPFILCSLMWSWVEQLKEPVIT
LAAKALANLNEXVEKEELKRKKELNSRDGAWERICGERDPFILCSLMWSWVEQLKEPVIT
LAAKALANLNEXVEKEELKRKKELNSRDGAWERICGERDPFILCSLMWSWVEQLKEPVIT

KEDVDMLVDRRADAAEALFLLEKGQHQTILCVLHCIVNLQTIPVDVEEAFLAHAIKAFTK
KEDVDMLVDRRADAAEALFLLEKGQHQTILCVLHCIVNLQTIPVDVEEAFLAHAIKAFTK
KEDVDMLVDRRADAAEALFLLEKGQHQTILCVLHCIVNLQTIPVDVEEAFLAHAIKAFTK

VNFDSENGPTVYNTLKKIFKHT       682
VNFDSENGP.VYNTLKKIFKHT
VNFDSENGPIVYNTLKKIFKHT       682
```

Fig. 5

HMMPFAM - alignment of 406 against pfam|hmm|DSPc

```
Dual specificity phosphatase, catalytic -
This hit is scoring at : -2.9
Scoring matrix : BLOSUM62 (used to infer consensus pattern)
Q:      38 NILAMARPsSELLEKYHII----DQFLSHGIKTIINLQRpgehaSCGNPLEQE-------
            .  .:.      L .    .       . . GI. :IN:..      .. NP.E :
H:       1 gpseilph.lYLGsystaseanlallsklgIthviNvte.....evpnpfeldkndrhyt

           -------SGFTYLPEA-FME-AgIYFYNFGWKDYGVASLTTILDMVKvmTFALQEGKVAI
                  SGFTYL. . . : . .Y.. :. :.    . ....:D.:.  . . : GKV.:
           nayisknsgftylqiPnvdDhI.YyhiawnhetdiskyfdeavdFId..karqkggkVLV

           HCHAGLGRTGVLIACYLVFATRMTADQAIIFV-----RAKRPNSIQTR-GQLLCVREFTQ
           HC.AG:.R:..LI..YL: ..:.:.::A. FV     :.:R   I... G :  : E:.:
           HCqAGiSRSAtliiAYLMktrnlslneAysfvyvYhikerRcpiisPNfgFmrQLieyer

           F      192

           k      172
```

SEQUENCE LISTING

<110> Bayer AG

<120> REGULATION OF HUMAN PTPASE-LIKE PROTIEN

<130> LIO382 Foreign Countries

<150> US 60/313,810

<151> 2001-08-22

<160> 14

<170> PatentIn version 3.1

<210> 1

<211> 2046

<212> DNA

<213> Homo sapiens

<400> 1
atggcatgtt ccatggcgtg tggcggtaga gcttgcaagt atgagaaccc agcccgctgg 60

agtgagcagg agcaagccat taagggggtt tactcatcct gggtcactga taatatactg 120

gccatggccc gcccatcctc tgagctcctg gagaagtacc acatcattga tcagttcctc 180

agccatggca taaaaacaat aatcaacctc cagcgccctg gtgagcatgc tagctgtggg 240

aaccctctgg aacaagaaag tggcttcaca taccttcctg aggctttcat ggaggctggc 300

atttacttct acaatttcgg atggaaggat tatggtgtag cgtctcttac tactatccta 360

gatatggtga aggtgatgac atttgcctta caggaaggaa aagtagctat ccattgtcat 420

gcagggcttg gtcgaacagg tgttttaata gcctgttact tagtttttgc aacgagaatg 480

actgctgacc aagcaattat atttgtgcgg gcaaagcgac ccaattccat acaaaccaga 540

ggacagctcc tctgtgtaag ggaatttact cagtttctaa ctcctctccg caatatattc 600

tcttgctgtg atcccaaagc acatgctgtc accttacctc aatatctaat tcgccagcgt 660

catctgcttc atggttatga ggcacgactt ctgaaacacg tgccaaaaat tatccaccta 720

gtttgcaaat tgctgctgga cttagcggag aacaggccag tgatgatgaa ggatgtgtcc 780

gaaggacctg gtctctctgc tgaaatagaa aagacaatgt ctgagatggt caccatgcag 840

ctggataaag agttactgag gcatgacagt gatgtgtcca acccgcctaa ccccactgca 900

gtggcagcag attttgacaa tcgaggcatg attttctcca atgagcaaca gtttgaccct 960

ctttggaaaa ggcggaatgt tgagtgcctt caacccctga ctcatctgaa aaggcggctc 1020

agctacagtg actcagattt aaagagggcc gagaacctcc tggagcaagg ggagactcca 1080

cagacagtgc ctgcccagat cttggttggc cacaagccca ggcagcagaa gctcataagc 1140

cattgttaca tcccacagtc tccagaacca gacttacaca aggaagcctt ggttcgcagc 1200

acactttctt tctggagtca gtcaaagttt ggaggcctgg aaggactcaa agataatggg 1260

tcaccaattt tccatggaag gatcattcca aaggaagcac agcagagtgg agctttctct 1320

gcagatgttt caggctcaca cagccctggg gagccagttt cacccagctt tgcaaatgtc 1380

cataaggatc caaaccctgc tcaccagcaa gtgtctcact gtcagtgtaa aactcatggt 1440

gttgggagcc ctggctctgt caggcagaac agcaggacac cccgaagccc tctggactgt 1500

ggctccagtc ccaaagcaca gttcttggtt gaacatgaaa cccaggacag taaagatctg 1560

tctgaagcag cttcacactc tgcattacag tctgaattga gtgctgaggc aagaagaata 1620

ctggcggcca aagccctagc aaatttaaat gaatctgtag aaaaggagga actaaaaagg 1680

aagaaagagc ttaattcccg agatggagct tgggaaagaa tatgtggcga gagggacccct 1740

ttcatcctat gcagcttgat gtggtcttgg gtggagcaac tgaaggagcc tgtaatcacc 1800

aaagaggatg tggacatgtt ggttgacagg cgagcagatg ccgcagaagc actttttta 1860

ttagagaagg gacagcacca gactattctc tgcgtgttgc actgcatagt gaacctgcag 1920

acaattcccg tggatgtgga ggaagctttc cttgcccatg ccattaaggc attcactaag 1980

gttaattttg attctgaaaa tggaccaaca gtttacaaca ccctgaagaa aatatttaag 2040

cacacg 2046

<210> 2

<211> 682

<212> PRT

<213> Homo sapiens

<400> 2

Met Ala Cys Ser Met Ala Cys Gly Gly Arg Ala Cys Lys Tyr Glu Asn
1 5 10 15

Pro Ala Arg Trp Ser Glu Gln Glu Gln Ala Ile Lys Gly Val Tyr Ser
20 25 30

Ser Trp Val Thr Asp Asn Ile Leu Ala Met Ala Arg Pro Ser Ser Glu
35 40 45

Leu Leu Glu Lys Tyr His Ile Ile Asp Gln Phe Leu Ser His Gly Ile
50 55 60

Lys Thr Ile Ile Asn Leu Gln Arg Pro Gly Glu His Ala Ser Cys Gly
65 70 75 80

Asn Pro Leu Glu Gln Glu Ser Gly Phe Thr Tyr Leu Pro Glu Ala Phe
85 90 95

Met Glu Ala Gly Ile Tyr Phe Tyr Asn Phe Gly Trp Lys Asp Tyr Gly
100 105 110

Val Ala Ser Leu Thr Thr Ile Leu Asp Met Val Lys Val Met Thr Phe
115 120 125

Ala Leu Gln Glu Gly Lys Val Ala Ile His Cys His Ala Gly Leu Gly
130 135 140

Arg Thr Gly Val Leu Ile Ala Cys Tyr Leu Val Phe Ala Thr Arg Met
145 150 155 160

Thr Ala Asp Gln Ala Ile Ile Phe Val Arg Ala Lys Arg Pro Asn Ser
165 170 175

Ile Gln Thr Arg Gly Gln Leu Leu Cys Val Arg Glu Phe Thr Gln Phe
180 185 190

Leu Thr Pro Leu Arg Asn Ile Phe Ser Cys Cys Asp Pro Lys Ala His
195 200 205

Ala Val Thr Leu Pro Gln Tyr Leu Ile Arg Gln Arg His Leu Leu His
210 215 220

Gly Tyr Glu Ala Arg Leu Leu Lys His Val Pro Lys Ile Ile His Leu
225 230 235 240

Val Cys Lys Leu Leu Leu Asp Leu Ala Glu Asn Arg Pro Val Met Met
245 250 255

Lys Asp Val Ser Glu Gly Pro Gly Leu Ser Ala Glu Ile Glu Lys Thr
260 265 270

Met Ser Glu Met Val Thr Met Gln Leu Asp Lys Glu Leu Leu Arg His
275 280 285

Asp Ser Asp Val Ser Asn Pro Pro Asn Pro Thr Ala Val Ala Ala Asp
290 295 300

Phe Asp Asn Arg Gly Met Ile Phe Ser Asn Glu Gln Gln Phe Asp Pro
305 310 315 320

Leu Trp Lys Arg Arg Asn Val Glu Cys Leu Gln Pro Leu Thr His Leu
325 330 335

Lys Arg Arg Leu Ser Tyr Ser Asp Ser Asp Leu Lys Arg Ala Glu Asn
340 345 350

Leu Leu Glu Gln Gly Glu Thr Pro Gln Thr Val Pro Ala Gln Ile Leu
355 360 365

Val Gly His Lys Pro Arg Gln Gln Lys Leu Ile Ser His Cys Tyr Ile
370 375 380

Pro Gln Ser Pro Glu Pro Asp Leu His Lys Glu Ala Leu Val Arg Ser
385 390 395 400

Thr Leu Ser Phe Trp Ser Gln Ser Lys Phe Gly Gly Leu Glu Gly Leu
405 410 415

Lys Asp Asn Gly Ser Pro Ile Phe His Gly Arg Ile Ile Pro Lys Glu
420 425 430

Ala Gln Gln Ser Gly Ala Phe Ser Ala Asp Val Ser Gly Ser His Ser
435 440 445

Pro Gly Glu Pro Val Ser Pro Ser Phe Ala Asn Val His Lys Asp Pro
450 455 460

```
Asn Pro Ala His Gln Gln Val Ser His Cys Gln Cys Lys Thr His Gly
465                 470                 475                 480

Val Gly Ser Pro Gly Ser Val Arg Gln Asn Ser Arg Thr Pro Arg Ser
                485                 490                 495

Pro Leu Asp Cys Gly Ser Ser Pro Lys Ala Gln Phe Leu Val Glu His
            500                 505                 510

Glu Thr Gln Asp Ser Lys Asp Leu Ser Glu Ala Ala Ser His Ser Ala
        515                 520                 525

Leu Gln Ser Glu Leu Ser Ala Glu Ala Arg Arg Ile Leu Ala Ala Lys
    530                 535                 540

Ala Leu Ala Asn Leu Asn Glu Ser Val Glu Lys Glu Glu Leu Lys Arg
545                 550                 555                 560

Lys Lys Glu Leu Asn Ser Arg Asp Gly Ala Trp Glu Arg Ile Cys Gly
                565                 570                 575

Glu Arg Asp Pro Phe Ile Leu Cys Ser Leu Met Trp Ser Trp Val Glu
            580                 585                 590

Gln Leu Lys Glu Pro Val Ile Thr Lys Glu Asp Val Asp Met Leu Val
        595                 600                 605

Asp Arg Arg Ala Asp Ala Ala Glu Ala Leu Phe Leu Leu Glu Lys Gly
    610                 615                 620

Gln His Gln Thr Ile Leu Cys Val Leu His Cys Ile Val Asn Leu Gln
625                 630                 635                 640

Thr Ile Pro Val Asp Val Glu Glu Ala Phe Leu Ala His Ala Ile Lys
                645                 650                 655

Ala Phe Thr Lys Val Asn Phe Asp Ser Glu Asn Gly Pro Thr Val Tyr
            660                 665                 670

Asn Thr Leu Lys Lys Ile Phe Lys His Thr
        675                 680
```

<210> 3

<211> 2086

<212> DNA

<213> macaca fascicularis

<220>

<221> misc_feature

<222> (1689)..(1689)

<223> unknown nucleotide

<300>

<308> EMBL/AB046040

<309> 2000-07-18

<400> 3
aggagagcgt ttacggcatg tcattcctgg acacatggcg tgttccatgg cgtgtggtgg 60

tagagcttgc aagtatgaga atccagcccg ctggagtgag caggaacaag ccattaaagg 120

ggtttactca tcctgggtca ctgataatat actggccatg gcccgcccat cctctgagct 180

cctggagaag taccacatca ttgatcagtt cctcagccat ggcataaaaa caataatcaa 240

cctccagcgc cctggtgagc atgctagctg tgggaaccct ctggaacaag aaagtggctt 300

cacatacctt cctgaggctt tcatggaggc tggcatttat ttctacaatt tcgggtggaa 360

ggattatggt gtagcatctc ttactactat cctagatatg gtgaaggtga tgacatttgc 420

cttacaggaa ggaaaagtag ctatccattg tcatgcaggg cttggtcgaa caggtgtctt 480

aatagcatgt tacttagttt ttgcaacaag aatgactgct gaccaagcaa ttatatttgt 540

gcgggcaaag cgacccaatt ccatacaaac cagagggcag ctcctctgtg taagggaatt 600

tactcagttt ctaactcctc tccgcaatat attctcttgc tgtgatccca aagcacatgc 660

tgtcacctta gctcaatatc taattcgcca gcgtcatctg cttcatggtt atgaggcacg 720

acttctgaaa catgtgccaa aaattatcca cctagtttgc aaattgcttc tggacttagc 780

ggaaaacagg ccagtgatga tgaaggacga gtccgaagga cctggtctct ctgctgaaat 840

agaaaagacc atgtctgaga tggtcaccat gcagctggat aaagaattac tgaggaatga 900

```
cagtgatgtg tccaacccgc ctaaccccac tgcagtggca gcagattttg acaatcgagg     960
catgattttc tccaatgagc aagagtttga ccctctttgg aaaaggcgga atgttgagtg    1020
ccttcaaccc ctttctcatc tgaaaaggcg gctcagctac agtgactcag atttaaagag    1080
ggccgagaac ctcctggagg aaggggagcc tccacagaca gtgcctgtcc aggtcttggt    1140
tggccacaac cccaggcagc agaagcgcat aagccattgt tacatcccac agtcttcaga    1200
accagactta cataaagaag ccttggttcg cagcacactt tctttctgga gtcagtctaa    1260
gtttggaagg ctggaagggc tcaaagataa tggctcacca attttccatg gacggatcat    1320
tccaaaggaa gcacagcaga gtagagcttt ctctgcagat gtttcaggcc cacacagccc    1380
tggggaacca gtttcaccca actttgcaaa tgtccgtaag gatccaaacc ctactcacca    1440
gcaagtgtct cactgtcagt gtaaaactca tggtgttggg agccctgggt ctgtcaggca    1500
gaacagcagg acaccccaaa gacctctgga ctgtggctgc agtctgaaag cacagttctt    1560
ggttgaacat gaaacccagg acagtaaaga tctggctgaa gcagcttcac actctacatt    1620
acagtctgaa ttgagtgctg aggcaagaag aatactggcg gccaaagctc tagcaaattt    1680
aaatgaatnt gtagaaaagg aggaactaaa aaggaagaaa gagcttaatt cccgagatgg    1740
agcttgggaa agaatatgtg gcgagaggga ccctttcatc ctatgcagct tgatgtggtc    1800
ttgggtggag caactgaagg agcctgtaat caccaaagag gatgtggaca tgttggttga    1860
caggcgagca gacgccgcag aagcactgtt tttattagag aagggacagc accagactat    1920
tctctgcgtg ttgcactgca tagtgaacct gcagacaatt cccgtggatg tggaggaagc    1980
tttccttgcc catgccatta aggcgttcac taaggttaat tttgattctg aaaatggacc    2040
aatagtttac aacaccctga agaaaatatt taagcacacg cggtga                   2086
```

<210> 4

<211> 380

<212> DNA

<213> Homo sapiens

<300>

<308> GenBank/AA907041

<309> 1998-04-13

<400> 4
cttccagcgt gtgcttaaat attttcttca gggtgttgta aactgttggt ccattttcag 60

aatcaaaatt aaccttagtg aatgccttaa tggcatgggc aaggaaagct tcctccacat 120

ccacgggaat tgtctgcagg ttcactatgc agtgcaacac gcagagaata gtctggtgct 180

gtcccttctc taataaaaaa agtgcttctg cggcatctgc tcgcctgtca accaacatgt 240

ccacatcctc tttggtgatt acaggctcct tcagttgctc cacccaagac cacatcaagc 300

tgcataggat gaaagggtcc ctctcgccac atattctttc ccaagctcca tctcgggaat 360

taagctcttt ctgccacatt 380

<210> 5

<211> 424

<212> DNA

<213> Homo sapiens

<300>

<308> GenBank/AI023460

<309> 1998-06-18

<400> 5
cttccagcgt gtgcttaaat attttcttca gggtgttgta aactgttggt ccattttcag 60

aatcaaaatt aaccttagtg aatgccttaa tggcatgggc aaggaaagct tcctccacat 120

ccacgggaat tgtctgcagg ttcactatgc agtgcaacac gcagagaata gtctggtgct 180

gtcccttctc taataaaaaa agtgcttctg cggcatctgc tcgcctgtca accaacatgt 240

ccacatcctc tttggtgatt acaggctcct tcagttgctc cacccaagac cacatcaagc 300

tgcataggat gaaagggtcc ctctcgccac atattctttc ccaagctcca tctcgggaat 360

taagctcttt ctaaaaagga agagaaagag acagacagac agagacagag acagggagag 420

gaga 424

<210> 6

<211> 1091

<212> DNA

<213> Homo sapiens

<300>

<308> GenBank/BG116082

<309> 2001-01-29

<400> 6
aaaaggagga actaaaaagg aaggtagaaa tgtggcagaa agagcttaat tcccgagatg 60

gagcttggga aagaatatgt ggcgagaggg accctttcat cctatgcagc ttgatgtggt 120

cttgggtgga gcaactgaag gagcctgtaa tcaccaaaga ggatgtggac atgttggttg 180

acaggcgagc agatgccgca gaagcacttt ttttattaga gaagggacag caccagacta 240

ttctctgcgt gttgcactgc atagtgaacc tgcagacaat tcccgtggat gtggaggaag 300

ctttccttgc ccatgccatt aaggcattca ctaaggttaa ttttgattct gaaaatggac 360

caacagttta caacaccctg aagaaaatat ttaagcacac gctggaagaa aaaagaaaaa 420

tgacaaaaga tggccctaag cctggcctct agctttcact cgtggtgaat atttcagacc 480

taaagatcca gatagtatct ctgttcatat gtgaataagt tgaagattgt ggggctactt 540

tttctcatag cactttatat gaatgtggta agtagtgctg agaatggtcg tccgtatttg 600

aaccaagtta ttattttaaa catatattaa ggctacattt aggtttgaaa aagtgccaat 660

aagttgtgtg ccactaacat tgattcatga gaactggagt gaaaatcatt ggcatgacca 720

tattaaagac aacggcggaa acagtctcat acataggaca tttggaaaca aacacacctg 780

ctcccaaata gccatctgga cggaaccaga caaaagaggt cgccaagagg aacacaagaa 840

gaggacccca gggggaaca gcccacgcaa agaactagag cgaacacgtg tgagagggag 900

aaaatcggga gcaaagcgga gtaagagaag aagaggagga aaacagacga aaggcaccgg 960

aaaagagagg gagaaggtga gagcaaacga aaagaggagg agagcgagca gagaagagag 1020

cagaagcaac agaagaggag aagagagaga gctaacaaga aagcacgcgt tgtgtcgctg 1080

tagtaagagt g 1091

<210> 7

<211> 524

<212> DNA

<213> Homo sapiens

<300>

<308> GenBank/AI829572

<309> 1999-07-12

<400> 7
ttccagcggg tgcttaaata ttttcttcag ggtgttgtaa actgttggtc cattttcaga 60

atcaaaatta accttagtga atgccttaat ggcatgggca aggaaagctt cctccacatc 120

cacgggaatt gtctgcaggt tcactatgca gtgcaacacg cagagaatag tctggtgctg 180

tcccttctct aataaaaaaa gtgcttctgc ggcatctgct cgcctgtcaa ccaacatgtc 240

cacatcctct ttggtgatta caggctcctt cagttgctcc acccaagacc acatcaagct 300

gcataggatg aaagggtccc tctcgccaca tattctttcc caagctccat ctcgggaatt 360

aagctctttc tgccacattt ctaccttcct ttttagttcc tccttttcta cagattcatt 420

taaatttgct agggctttgg ccgccagtat tcttcttgcc tcagcactca attcagactg 480

taatgcagag tgtgaagctg cttcagacag atctttactg tcct 524

<210> 8

<211> 296

<212> DNA

<213> Homo sapiens

<300>

<308> GenBank/AA910424

<309> 1998-04-13

<400> 8
ttttcttcag ggtgttgtaa actgttggtc cattttcaga atcaaaatta accttagtga 60

atgccttaat ggcatgggca aggaaagctt cctccacatc cacgggaatt gtctgcaggt 120

tcactatgca gtgcaacacg cagagaatag tctggtgctg tcccttctct aataaaaaaa 180

gtgcttctgc ggcatctgct cgcctgtcaa ccaacatgtc cacatcctct ttggtgatta 240

caggctcctt cagttgctcc acccaagacc acatcaagct gcataggatg aaaggg 296

<210> 9

<211> 459

<212> DNA

<213> Homo sapiens

<300>

<308> GenBank/AA651920

<309> 1997-11-03

<400> 9
tttttttttt tttttttttt ttaatttttc aaaacaaaaa tgtagcttaa atatatttta 60

aaataaataa ttggttcaaa tacggacgac cattctcagc acaaactaac aacattcaaa 120

ataaagtgct atgagaaaaa gtagccccac aatcttcaac ttattcacat atgaacagag 180

atactatctg gatctttagg tctgaaatat tcaccacgag tgaaagctag aggccaggct 240

tagggccatc ttttgtcatt tttctttttt cttccagcgt gtgcttaaat attttcttca 300

gggtgttgta aactgttggt ccattttcag aatcaaaatt aaccttagtg aatgccttaa 360

tggcatgggc aaggaaagct tcctccacat ccacgggaat tgtctgcagg ttcactatgc 420

agtgcaacac gcagagaata gtctggtgct gtcccttct 459

<210> 10

<211> 610

<212> DNA

<213> Homo sapiens

<300>

<308> GenBank/BE897818

<309> 2000-09-26

<400> 10
gtgagcagag gatcggtgcc gcgccgtgag ggtcgtagag gctgcctcgt ggcccgctat 60

ccacagaccc tccgcggggc cagcgtctcc gggccggccg cgcctggtcc gcgaccgcag 120
acgacttgtg acacggacgg cgcgtctgct gaccccagcc gggccgaggg cttggaatct 180
cgagcgggaa aagactaaca tggctgcagg agtcttgcct cagaatgaac aaccatattc 240
taccttggtg aataacagcg agtgtgttgc aaaatgaaag tttacttcta caatttcgga 300
tggaaggatt atggtgtagc gtctcttact actatcctag atatggtgaa gggtgatgac 360
atttgcttac aggaacggac caagtagcta tccattgtca atgccagggc acttgggtcg 420
caacacggtg tttctacata cgcctgttac tctagatatt tgcaccgaga aatgacctgc 480
ctgacccaac gccaattata ttctgtgccg ggccaaagcg accaattcca ctacaaaacc 540
acaggacagc tccctctgtg taagggactt tactcacgtt ctaactcctc tccgcaccta 600
tcctttgtgg 610

<210> 11
<211> 319
<212> DNA
<213> Homo sapiens

<300>
<308> GenBank/Z44764
<309> 1994-11-14

<400> 11
gaaaatggac caacagttta caacaccctg aagaaaatat ttaagcacac gctggaagaa 60
aaaagaaaaa tgacaaaaga tggccctaag cctggcctct agctttcact cgtggtgaat 120
atttcagacc taaagatcca gatagtatct ctgttcatat gtgaataagt tgaagattgt 180
ggggctactt tttctcatag cactttattt tgaatgttgt tagtttgtgc tgagaatggt 240
cgtccgtatt tgaaccaatt atttatttta aaatatattt aagctacatt tttgttttga 300
aaaattgcca taaatttgg 319

<210> 12
<211> 24

<212> DNA

<213> Artificial sequence

<220>

<223> random primer

<400> 12
tcaactgact agatgtacat ggac 24

<210> 13

<211> 7

<212> PRT

<213> Artificial sequence

<220>

<223> consensus motif

<220>

<221> MISC_FEATURE

<222> (3)..(4)

<223> any amino acid

<400> 13

Lys Cys Xaa Xaa Tyr Trp Pro
1 5

<210> 14

<211> 10

<212> PRT

<213> Artificial sequence

<220>

<223> consensus motif

<220>

<221> MISC_FEATURE

<222> (4)..(4)

<223> any amino acid

<220>

<221> MISC_FEATURE

<222> (6)..(6)

<223> any amino acid

<220>

<221> MISC_FEATURE

<222> (9)..(9)

<223> any amino acid

<400> 14

```
His Cys Ser Xaa Gly Xaa Gly Arg Xaa Gly
1               5                   10
```

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 02/09331

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/52 C12N9/00 C12N9/02 C07K14/37 C12N15/82
C12N1/21 C12N15/53 C12N15/31 G01N33/573 G01N33/68
A61K35/00 A01N61/00 A01H5/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, EMBL, BIOSIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL 'Online! 18 July 2000 (2000-07-18) HASHIMOTO K. ET AL.: "Macaca fascicularis brain cDNA, clone :QccE-10114" retrieved from EMBL Database accession no. AB046040 XP002225534 abstract -& DATABASE TREMBL 'Online! 15 October 2000 (2000-10-15) "Unnamed protein product" retrieved from TREMBL Database accession no. Q9N091 XP002225535 abstract	1-12
P,X	-& OSADA N ET AL: "Assignement of 118 novel cDNAs of cynomolgus monkey brain to human chromosomes" GENE: AN INTERNATIONAL JOURNAL ON GENES -/--	

[X] Further documents are listed in the continuation of box C.

[ ] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
14 January 2003	28/01/2003

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Turri, M

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
	AND GENOMES, ELSEVIER SCIENCE PUBLISHERS, BARKING, GB, vol. 275, no. 1, 5 September 2001 (2001-09-05), pages 31-37, XP004307108 ISSN: 0378-1119	
X	FISCHER E H ET AL: "PROTEIN TYROSINE PHOSPHATASES: A DIVERSE FAMILY OF INTRACELLULAR AND TRANSMEMBRANE ENZYMES" SCIENCE, AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE,, US, vol. 253, 26 July 1991 (1991-07-26), pages 401-406, XP001026247 ISSN: 0036-8075 the whole document See especially page 402, left-hand column, first paragraph	1-12
A	YI T ET AL: "IDENTIFICATION OF NOVEL PROTEIN TYROSINE PHOSPHATASES OF HEMATOPOIETIC CELLS BY POLYMERASE CHAIN REACTION AMPLIFICATION" BLOOD, W.B. SAUNDERS, PHILADELPHIA, VA, US, vol. 78, no. 9, 1 November 1991 (1991-11-01), pages 2222-2228, XP002034263 ISSN: 0006-4971 page 2223; figure 1	
A	KRUEGER N X ET AL: "STRUCTURAL DIVERSITY AND EVOLUTION OF HUMAN RECEPTOR-LIKE PROTEIN TYROSINE PHASPHATASES" EMBO JOURNAL, OXFORD UNIVERSITY PRESS, SURREY, GB, vol. 9, no. 10, 1 October 1990 (1990-10-01), pages 3241-3252, XP002032773 ISSN: 0261-4189 cited in the application	
A	YI T ET AL: "PROTEIN TYROSINE PHOSPHATASE CONTAINING SH2 DOMAINS CHARACTERIZATION PREFERENTIAL EXPRESSION IN HEMATOPOIETIC CELLS AND LOCALIZATION TO HUMAN CHROMOSOME 12P12-P13" MOLECULAR AND CELLULAR BIOLOGY, vol. 12, no. 2, 1992, pages 836-846, XP009003191 ISSN: 0270-7306 cited in the application	

-/--

International Application No
PCT/EP 02/09331

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	DATABASE PROSITE 'Online! November 1990 (1990-11) "TYR_phosphatase_1; PATTERN" retrieved from PROSITE Database accession no. PS00383 XP002225536	
A	-& DATABASE PROSITE 'Online! July 1999 (1999-07) "Tyrosine specific protein phosphatesase signature and profiles" retrieved from PROSITE Database accession no. PDOC00323 XP002225537 -----	

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 13-17

Present claims 13-17 relate to a compound defined by reference to a desirable characteristic or property, namely reducing or modulating the activity of the PTPase-like protein.

The claims cover all compounds having this characteristic or property, whereas the application does not provide support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for any of them. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the compound by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search over the claimed scope impossible.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

International application No.
PCT/EP 02/09331

INTERNATIONAL SEARCH REPORT

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: 13-17
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.